

OM of: US-09-397-342-1 to: EST:* out_format: pfs

Date: Apr 28, 2001 7:13 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+g2n.model -DEV=xlp
-O/cgn2.1/SPFO.spool/US09397342/runat_27042001.165959.22879/app.query.fasta.1.384
-DB=EST -QEXT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
-MINNAATCH=0.050 -LOOPEXT=0.000 -LOOPEXT=0.500 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FCGAPOP=6.000
-FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DEL0P=6.000
-DEL0EXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09397342@cgn1.1.2985 -NCPU=6
-ICPU=3 -LONGLOGS -NO_XLPHY -WAIT -THREADS=1
```

Arch information block:

```
Query: US-09-397-342-1
Query length: 323
Database: EST:*
Database sequences: 9623517
Database length: 73081774
Search time (sec): 1149.400000
```

Sequence list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est42:AM157357	-	831.00	1689.59	5.1e-85	483	AM157357.a88a06.x1.Schneider
gb_est42:AM157486	-	831.00	1689.59	5.1e-85	483	AM157486.a88a06.x1.Schneider
gb_est42:AM157496	-	831.00	1687.29	6.8e-85	483	AM157496.a88a06.x1.Schneider
gb_est42:AM157499	-	831.00	1554.74	1.7e-77	633	AM157499.a88a06.x1.Schneider
gb_est42:AM157500	-	831.00	1554.74	1.7e-77	633	AM157500.a88a06.x1.Schneider
gb_est42:AM157501	-	831.00	1554.74	1.7e-77	633	AM157501.a88a06.x1.Schneider
gb_est42:AM157502	-	831.00	1554.74	1.7e-77	633	AM157502.a88a06.x1.Schneider
gb_est42:AM157503	-	831.00	1554.74	1.7e-77	633	AM157503.a88a06.x1.Schneider
gb_est42:AM157504	-	831.00	1554.74	1.7e-77	633	AM157504.a88a06.x1.Schneider
gb_est42:AM157505	-	831.00	1554.74	1.7e-77	633	AM157505.a88a06.x1.Schneider
gb_est42:AM157506	-	831.00	1554.74	1.7e-77	633	AM157506.a88a06.x1.Schneider
gb_est42:AM157507	-	831.00	1554.74	1.7e-77	633	AM157507.a88a06.x1.Schneider
gb_est42:AM157508	-	831.00	1554.74	1.7e-77	633	AM157508.a88a06.x1.Schneider
gb_est42:AM157509	-	831.00	1554.74	1.7e-77	633	AM157509.a88a06.x1.Schneider
gb_est42:AM157510	-	831.00	1554.74	1.7e-77	633	AM157510.a88a06.x1.Schneider
gb_est42:AM157511	-	831.00	1554.74	1.7e-77	633	AM157511.a88a06.x1.Schneider
gb_est42:AM157512	-	831.00	1554.74	1.7e-77	633	AM157512.a88a06.x1.Schneider
gb_est42:AM157513	-	831.00	1554.74	1.7e-77	633	AM157513.a88a06.x1.Schneider
gb_est42:AM157514	-	831.00	1554.74	1.7e-77	633	AM157514.a88a06.x1.Schneider
gb_est42:AM157515	-	831.00	1554.74	1.7e-77	633	AM157515.a88a06.x1.Schneider
gb_est42:AM157516	-	831.00	1554.74	1.7e-77	633	AM157516.a88a06.x1.Schneider
gb_est42:AM157517	-	831.00	1554.74	1.7e-77	633	AM157517.a88a06.x1.Schneider
gb_est42:AM157518	-	831.00	1554.74	1.7e-77	633	AM157518.a88a06.x1.Schneider
gb_est42:AM157519	-	831.00	1554.74	1.7e-77	633	AM157519.a88a06.x1.Schneider
gb_est42:AM157520	-	831.00	1554.74	1.7e-77	633	AM157520.a88a06.x1.Schneider
gb_est42:AM157521	-	831.00	1554.74	1.7e-77	633	AM157521.a88a06.x1.Schneider
gb_est42:AM157522	-	831.00	1554.74	1.7e-77	633	AM157522.a88a06.x1.Schneider
gb_est42:AM157523	-	831.00	1554.74	1.7e-77	633	AM157523.a88a06.x1.Schneider
gb_est42:AM157524	-	831.00	1554.74	1.7e-77	633	AM157524.a88a06.x1.Schneider
gb_est42:AM157525	-	831.00	1554.74	1.7e-77	633	AM157525.a88a06.x1.Schneider
gb_est42:AM157526	-	831.00	1554.74	1.7e-77	633	AM157526.a88a06.x1.Schneider
gb_est42:AM157527	-	831.00	1554.74	1.7e-77	633	AM157527.a88a06.x1.Schneider
gb_est42:AM157528	-	831.00	1554.74	1.7e-77	633	AM157528.a88a06.x1.Schneider
gb_est42:AM157529	-	831.00	1554.74	1.7e-77	633	AM157529.a88a06.x1.Schneider
gb_est42:AM157530	-	831.00	1554.74	1.7e-77	633	AM157530.a88a06.x1.Schneider
gb_est42:AM157531	-	831.00	1554.74	1.7e-77	633	AM157531.a88a06.x1.Schneider
gb_est42:AM157532	-	831.00	1554.74	1.7e-77	633	AM157532.a88a06.x1.Schneider
gb_est42:AM157533	-	831.00	1554.74	1.7e-77	633	AM157533.a88a06.x1.Schneider
gb_est42:AM157534	-	831.00	1554.74	1.7e-77	633	AM157534.a88a06.x1.Schneider
gb_est42:AM157535	-	831.00	1554.74	1.7e-77	633	AM157535.a88a06.x1.Schneider
gb_est42:AM157536	-	831.00	1554.74	1.7e-77	633	AM157536.a88a06.x1.Schneider
gb_est42:AM157537	-	831.00	1554.74	1.7e-77	633	AM157537.a88a06.x1.Schneider
gb_est42:AM157538	-	831.00	1554.74	1.7e-77	633	AM157538.a88a06.x1.Schneider
gb_est42:AM157539	-	831.00	1554.74	1.7e-77	633	AM157539.a88a06.x1.Schneider
gb_est42:AM157540	-	831.00	1554.74	1.7e-77	633	AM157540.a88a06.x1.Schneider
gb_est42:AM157541	-	831.00	1554.74	1.7e-77	633	AM157541.a88a06.x1.Schneider
gb_est42:AM157542	-	831.00	1554.74	1.7e-77	633	AM157542.a88a06.x1.Schneider
gb_est42:AM157543	-	831.00	1554.74	1.7e-77	633	AM157543.a88a06.x1.Schneider
gb_est42:AM157544	-	831.00	1554.74	1.7e-77	633	AM157544.a88a06.x1.Schneider
gb_est42:AM157545	-	831.00	1554.74	1.7e-77	633	AM157545.a88a06.x1.Schneider
gb_est42:AM157546	-	831.00	1554.74	1.7e-77	633	AM157546.a88a06.x1.Schneider
gb_est42:AM157547	-	831.00	1554.74	1.7e-77	633	AM157547.a88a06.x1.Schneider
gb_est42:AM157548	-	831.00	1554.74	1.7e-77	633	AM157548.a88a06.x1.Schneider
gb_est42:AM157549	-	831.00	1554.74	1.7e-77	633	AM157549.a88a06.x1.Schneider
gb_est42:AM157550	-	831.00	1554.74	1.7e-77	633	AM157550.a88a06.x1.Schneider
gb_est42:AM157551	-	831.00	1554.74	1.7e-77	633	AM157551.a88a06.x1.Schneider
gb_est42:AM157552	-	831.00	1554.74	1.7e-77	633	AM157552.a88a06.x1.Schneider
gb_est42:AM157553	-	831.00	1554.74	1.7e-77	633	AM157553.a88a06.x1.Schneider
gb_est42:AM157554	-	831.00	1554.74	1.7e-77	633	AM157554.a88a06.x1.Schneider
gb_est42:AM157555	-	831.00	1554.74	1.7e-77	633	AM157555.a88a06.x1.Schneider
gb_est42:AM157556	-	831.00	1554.74	1.7e-77	633	AM157556.a88a06.x1.Schneider
gb_est42:AM157557	-	831.00	1554.74	1.7e-77	633	AM157557.a88a06.x1.Schneider
gb_est42:AM157558	-	831.00	1554.74	1.7e-77	633	AM157558.a88a06.x1.Schneider
gb_est42:AM157559	-	831.00	1554.74	1.7e-77	633	AM157559.a88a06.x1.Schneider
gb_est42:AM157560	-	831.00	1554.74	1.7e-77	633	AM157560.a88a06.x1.Schneider
gb_est42:AM157561	-	831.00	1554.74	1.7e-77	633	AM157561.a88a06.x1.Schneider
gb_est42:AM157562	-	831.00	1554.74	1.7e-77	633	AM157562.a88a06.x1.Schneider
gb_est42:AM157563	-	831.00	1554.74	1.7e-77	633	AM157563.a88a06.x1.Schneider
gb_est42:AM157564	-	831.00	1554.74	1.7e-77	633	AM157564.a88a06.x1.Schneider
gb_est42:AM157565	-	831.00	1554.74	1.7e-77	633	AM157565.a88a06.x1.Schneider
gb_est42:AM157566	-	831.00	1554.74	1.7e-77	633	AM157566.a88a06.x1.Schneider
gb_est42:AM157567	-	831.00	1554.74	1.7e-77	633	AM157567.a88a06.x1.Schneider
gb_est42:AM157568	-	831.00	1554.74	1.7e-77	633	AM157568.a88a06.x1.Schneider
gb_est42:AM157569	-	831.00	1554.74	1.7e-77	633	AM157569.a88a06.x1.Schneider
gb_est42:AM157570	-	831.00	1554.74	1.7e-77	633	AM157570.a88a06.x1.Schneider
gb_est42:AM157571	-	831.00	1554.74	1.7e-77	633	AM157571.a88a06.x1.Schneider
gb_est42:AM157572	-	831.00	1554.74	1.7e-77	633	AM157572.a88a06.x1.Schneider
gb_est42:AM157573	-	831.00	1554.74	1.7e-77	633	AM157573.a88a06.x1.Schneider
gb_est42:AM157574	-	831.00	1554.74	1.7e-77	633	AM157574.a88a06.x1.Schneider
gb_est42:AM157575	-	831.00	1554.74	1.7e-77	633	AM157575.a88a06.x1.Schneider
gb_est42:AM157576	-	831.00	1554.74	1.7e-77	633	AM157576.a88a06.x1.Schneider
gb_est42:AM157577	-	831.00	1554.74	1.7e-77	633	AM157577.a88a06.x1.Schneider
gb_est42:AM157578	-	831.00	1554.74	1.7e-77	633	AM157578.a88a06.x1.Schneider
gb_est42:AM157579	-	831.00	1554.74	1.7e-77	633	AM157579.a88a06.x1.Schneider
gb_est42:AM157580	-	831.00	1554.74	1.7e-77	633	AM157580.a88a06.x1.Schneider
gb_est42:AM157581	-	831.00	1554.74	1.7e-77	633	AM157581.a88a06.x1.Schneider
gb_est42:AM157582	-	831.00	1554.74	1.7e-77	633	AM157582.a88a06.x1.Schneider
gb_est42:AM157583	-	831.00	1554.74	1.7e-77	633	AM157583.a88a06.x1.Schneider
gb_est42:AM157584	-	831.00	1554.74	1.7e-77	633	AM157584.a88a06.x1.Schneider
gb_est42:AM157585	-	831.00	1554.74	1.7e-77	633	AM157585.a88a06.x1.Schneider
gb_est42:AM157586	-	831.00	1554.74	1.7e-77	633	AM157586.a88a06.x1.Schneider
gb_est42:AM157587	-	831.00	1554.74	1.7e-77	633	AM157587.a88a06.x1.Schneider
gb_est42:AM157588	-	831.00	1554.74	1.7e-77	633	AM157588.a88a06.x1.Schneider
gb_est42:AM157589	-	831.00	1554.74	1.7e-77	633	AM157589.a88a06.x1.Schneider
gb_est42:AM157590	-	831.00	1554.74	1.7e-77	633	AM157590.a88a06.x1.Schneider
gb_est42:AM157591	-	831.00	1554.74	1.7e-77	633	AM157591.a88a06.x1.Schneider
gb_est42:AM157592	-	831.00	1554.74	1.7e-77	633	AM157592.a88a06.x1.Schneider
gb_est42:AM157593	-	831.00	1554.74	1.7e-77	633	AM157593.a88a06.x1.Schneider
gb_est42:AM157594	-	831.00	1554.74	1.7e-77	633	AM157594.a88a06.x1.Schneider
gb_est42:AM157595	-	831.00	1554.74	1.7e-77	633	AM157595.a88a06.x1.Schneider
gb_est42:AM157596	-	831.00	1554.74	1.7e-77	633	AM157596.a88a06.x1.Schneider
gb_est42:AM157597	-	831.00	1554.74	1.7e-77	633	AM157597.a88a06.x1.Schneider
gb_est42:AM157598	-	831.00	1554.74	1.7e-77	633	AM157598.a88a06.x1.Schneider
gb_est42:AM157599	-	831.00	1554.74	1.7e-77	633	AM157599.a88a06.x1.Schneider
gb_est42:AM157600	-	831.00	1554.74	1.7e-77	633	AM157600.a88a06.x1.Schneider
gb_est42:AM157601	-	831.00	1554.74	1.7e-77	633	AM157601.a88a06.x1.Schneider
gb_est42:AM157602	-	831.00	1554.74	1.7e-77	633	AM157602.a88a06.x1.Schneider
gb_est42:AM157603	-	831.00	1554.74	1.7e-77	633	AM157603.a88a06.x1.Schneider
gb_est42:AM157604	-	831.00	1554.74	1.7e-77	633	AM157604.a88a06.x1.Schneider
gb_est42:AM157605	-	831.00	1554.74	1.7e-77	633	AM157605.a88a06.x1.Schneider
gb_est42:AM157606	-	831.00	1554.74	1.7e-77	633	AM1

151 ACCACTATGATACAGTGAACACTACTTGATTGATACACCACTTGA 200
 225 WAPASnlleMethrHISglyLeuSerSerLeuSerGlyLeuValA 242
 201 GGACAAATTCATGACTCAGCGTTTATCAAGTTTATGTTCTGACTGTA 250
 242 laseRleleuGlyThrProAlaaspVallellySerArgilleMetan 258
 251 CTTCTATCTCTGGAACACACCGCATGTCATCAAAAGCAGATATGAT 300
 259 GlnProArgaspLysGlnGlyArgGlyLeuLeuTyrLysSerSerThrAs 275
 301 CAACACGAGATAAACAAGAGGAGGACTTTGTATTAATCATCGACGTA 350
 275 pCysleuileGlnAlaValGlnGlyGlyGlyPheMetSerLeuTyrLysG 292
 351 CTGCTGATTCAGGCTGTTCAAGGTGAAGGATTCATGATCATATATAAG 400
 292 lypheLeuProSerTrpleuArgMet 300
 401 GCTTTTACCATCTTGCTGAGAAATG 426
 seq_name: gb_est42:AW161176
 seq_documentation_block: 564 bp mRNA EST 09-NOV-1999
 LOCUS AW161176 au79d09.y1 schneider fetal brain 00004 Homo sapiens cDNA clone
 DEFINITION IMAGE:2782481 5' similar to SW:0CP4_HUMAN 095847 MITOCHONDRIAL
 UNCOUPLING PROTEIN 4 ; mRNA sequence.
 ACCESSION AW161176
 VERSION AW161176.1 GI:6300209
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 564)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marris,M., Martin,
 J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
 White,Y., Wylie,T., Waterson,R. and Wilson,R.
 WashU-NCI human EST project
 Unpublished (1997)
 Other ESTs: au79d09.x1
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 435.
 FEATURES
 location/Qualifiers
 1..564
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2782481"
 /clone_lib="Schneider fetal brain 00004"
 /sex="male"
 /tissue_type="frontal lobe"
 /dev_stage="5 months post-conception"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript SK (Stratagene);
 Site:1: SstI; Site:2: XhoI; Double-stranded cDNA was
 prepared from human fetal brain tissue. 5' and 3'
 adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-GAAGAGAGAAGAGCTCAAGATCCTTAATTAATATCCCCCCCCC-3'
 and 3' adaptor sequence:
 5'-GAGAGAGAGCTCGAGTCTTTTCTTTTCTTTT-3'. The library was
 size-selected for >0.5 kb inserts and has an average

insert size estimated at 1.2 kb. This library was
 constructed using the CAP-trapper method for full-length
 enrichment and has not undergone amplification. Library
 was constructed by Dr. Claudio Schneider (UNCIB-Area
 Science Park, Trieste, Italy)."
 BASE COUNT 129 a 148 c 180 g 107 t
 ORIGIN
 alignment_scores:
 Quality: 676.00 Length: 130
 Ratio: 5.200 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-397-342-1 x AW161176 ..
 Align seg 1/1 to: AW161176 from: 1 to: 564
 1 MetSerValProGluGluGluAlaArgLeuLeuProLeuThrGlnArgTr 17
 175 ATGTCGTCGCCGAGAGAGAGAGAGAGGCTTTGGCCGTGACCCAGAGATG 224
 17 ProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
 225 GCCCGAGACGAGCAATTCCTACTGCTCGGCTGCCGGCTACCGCGCG 274
 34 luleAlaIaThrPheProLeuaspLeuThrLysThrArgLeuGlnMetGln 50
 275 AGCTGACCAACCTTCCCTCGATCTCACAAAACATCGACTCCAAATGCA 324
 51 GlyGluAlaAlaLeuAlaArgLeuGlyaspGlyAlaArgGluSerAlaArg 67
 325 GGAGAGAGAGCTCTTGCTGCTGGGTGGAGAGCGTGCAAGAGATTCGCC 374
 67 OTyArgGlyMetValArgThrAlaLeuGlylIleleGluGluGluGlyP 84
 375 CTATGGGGGATGTCGCCACACGCCCTAGGATCTTTAAAGAGAGAGCT 424
 84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValVal 100
 425 TTCTAAAGCTTTGGCAAGAGAGAGACACCGCATTTACAGACAGCTAGT 474
 101 TyrSerGlyGlyArgMetValThrTyrGlnHisLeuAlaGluValValph 117
 475 TATCTTGAGGTCGATGTCACATATGACATATCCCGAGAGGTTGT 524
 117 eGlyLysSerGluAspGluHisTyrProLeuTrpLysSer 130
 525 TGGCAAAAGTGAGATGAGCATTTATCCCTTTGGAAATCA 564
 seq_name: gb_est42:AW147976
 seq_documentation_block: 803 bp mRNA EST 22-JUN-2000
 LOCUS AW147976 da01d08.x1 xenopus laevis oocyte xenopus laevis cDNA clone
 DEFINITION XENOPUS_SOURCE_ID:xlnoc001b16 3 similar to SW:0CP4_HUMAN 095847
 MITOCHONDRIAL UNCOUPLING PROTEIN 4 ; mRNA sequence.
 ACCESSION AW147976
 VERSION AW147976.1 GI:6195872
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 803)
 AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
 B., Gibbons,M., Harvey,N., Rittler,E., Jackson,Y., McCann,R.,
 Waterson,R. and Wilson,R.
 WashU xenopus EST project, 1999
 TITLE Unpublished (1999)
 JOURNAL

size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy).

BASE COUNT 115 a 141 c 173 g 96 t 1 others

alignment_scores:

Quality: 589.00 Length: 115
Ratio: 5.122 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-397-342-1 x A1878851 ..

Align seg 1/1 to: A1878851 from: 1 to: 526

```

1 MetSerValProGluGluGluArgLeuLeuProLeuThrGluArgTr 17
|||||
175 ATGTCCCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 224
|||||
17 PProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
|||||
225 GCCCGGAGGAGCAAAATTCCTACTGTCGGCTGGCGGCTACCTGGCCG 274
|||||
34 IuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGluMetGln 50
|||||
275 AGCTAGCAACCTTCCCTGGATCTCACAACCAATCAGCTCAATGCA 324
|||||
51 GLVGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAlaPr 67
|||||
325 GGAGAGACAGCTCTTGGCTGGTGGAGACGCTGAGAGAGAGATCTGCC 374
|||||
67 OTYrArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGluGlyP 84
|||||
375 CTATAGGGGAGATGCTGGCAGACAGCCCTAGAGGATCATTTGAAGAGAAGCT 424
|||||
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTYrArgHisValAl 100
|||||
425 TTCTAAAGCTNTGGCAAGAGAGTACACCCGCAATTTACAGACAGTGTG 474
|||||
101 TYrSerGlyGlyArgMetValThrTYrGluHisLeuArgGluVal 115
|||||
475 TATCTGAGGTGCAATGTGCATATGACATCTCCGAGAGGTG 519
|||||

```

q_name: gb_est69:BE061032

seq_documentation_block:

LOCUS BE061032 607 bp mRNA EST 09-JUN-2000
DEFINITION QV0-BT0041-061099-033-c02 BT0041 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE061032
VERSION BE061032.1 GI:8405682

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 607)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL

MEDLINE

COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV0-BT0041-061
099-033-c02&ts=1999-10-06&ft=1)

Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 596.
Location/Qualifiers

FEATURES

source

1..607
/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="BT0041"

/dev_stage="Adult"

/note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
'716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 187 a 119 c 143 g 158 t

ORIGIN

alignment_scores:

Quality: 583.00 Length: 187
Ratio: 4.485 Gaps: 4
Percent Similarity: 69.519 Percent Identity: 66.310

alignment_block:

US-09-397-342-1 x BE061032 ..

Align seg 1/1 to: BE061032 from: 1 to: 607

```

95 IleTYrArgHisVal..... 99
|||||
164 ATACATAGCATATATAGATATAGCAAGTTACTTTTAATGATCT 213
|||||
100 .....ValTYrSerGlyGlyArgMetValThrTYrGluHisLeuA 113
|||||
214 TTTTACCAGGTGATTCCTGGAGGTGCAATGTGCATATGACATATCC 263
|||||
113 rGluValAlaPheGlyLysSerGluAspGluHisTYrProLeuTrpLys 129
|||||
264 GAGAGTTGTGTTGGCAAAAGTGAAGTGAAGCATATCCCTTTGGAAA 313
|||||
130 SerValIleGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAs 146
|||||
314 TCAGTCATTTGAGGAGATGATGCTGTGTATTTAGCCAGTTTACGCAA 363
|||||
146 nProThrAspLeuValLysValGlnMetGlnMetGluGlyLysArgLysL 163
|||||
364 TCCAACTGACCTAGTGAAGGTTCAGATGCCAAATGAGAAAGGAAGAAC 413
|||||
163 euGluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 179
|||||
414 TCGAAGGAAACCATTCGATTT..... 436
|||||
180 IleLeuAlaGluGlyGlyIleArgGlyLeuThrPalaGlyTrpValProAs 196
|||||
436 ..... 436
|||||
196 nIleGluArgAlaAlaLeuValAsnMetGlyAspLeuThrThrTYrAsp 212
|||||
437 .....AACCATTTATGAT 449
|||||
213 ThrValLysHisTYrLeuValIleuAsnThrProLeuGluAspAsnIleMe 229
|||||

```

```

|||||
450 ACAGTGAACACTACTTGGATTGATACACCACTTGAGACATATCAAT 499
229 tTtHrHsGLyLeuSerSerLeuCySerGlyLeuValAlaSerIleLeu 246
500 GACTCAGCGTTTATCAAGTTATGTTCTGAGCTGAGCTTCTATCTG 549
246 tTtHrProAlaAspValIleLysSerArgIleMetAsnGlnProArgasp 262
550 GAACACACACCGATGTCATCAAGACGAATATGAAAT...CCACACACC 596
263 LysGlnGly 265
597 AAGGAGGT 605
seq_name: gb_est199:FL11430

```

```

seq_documentation_block:
LOCUS   FL11430       341 bp      mRNA      EST      12-MAR-1995
DEFINITION   HSC2MH061 normalized infant brain cDNA Homo sapiens cDNA clone
C-2wh06, mRNA sequence.
ACCESSION   FL11430
VERSION     FL11430.1  GI:705728
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 341)
REFERENCE   1 (bases 1 to 341)
AUTHORS    Aufferay,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Houdigatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Mariage-Samson,R., Pletu,G., Pouliot,Y.,
SebastienI-Kabaktchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
JOURNAL    95277534
MEDLINE    Contact: Genethon
COMMENT    Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C; Genexpress_sequence_id: y1c-2wh06
Seq primer: (-21)M13 universal.

```

```

FEATURES
    source
        location/Qualifiers
            1..341
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="c-2wh06"
                /clone_lib="normalized infant brain cDNA"
                /sex="Female"
                /tissue_type="total brain"
                /dev_stage="3 months old"
                /note="Organ: brain; Vector: lafmid BA; Site:1: HindIII;
                Site:2: NotI; sex:Female; dev stage=3 months old;
                isolate=muscular atrophy patient; tissue_type=total brain
                ; total mRNA was oligo (dT) primed and directionally
                cloned 5' -> 3' into the HindIII -> NotI sites of the
                lafmid BA vector. Clone library from B.Souares, Psychiatry
                Dept. Columbia University, USA. Normalization_method:
                Bento Soares, P.N.A.S in press"
BASE COUNT      86 a      85 c      96 g      72 t      2 others
ORIGIN

```

```

alignment_scores:
Quality:      567.00      Length:      113
Ratio:        5.108      Gaps:      0
Percent Similarity: 98.230      Percent Identity: 98.230

```

```

alignment_block:
US-09-397-342-1 x FL11430
Align seg 1/1 to: FL11430 from: 1 to: 341

```

```

10 LeuLeuProLeuThrGlnArgTrpProArgAlaSerLysPheLeuLeu 26
3 CTTTTCCTCCCTGACCCAGATGGCCCCCGAGGAGCAAAATTCCTACTGTC 52
26 rGlyCysAlaAlaThrValAlaGluLeuAlaThrPheProLeuAspLeu 43
53 CGGCTCGCGGCTACCGCTGCGGAGCTACCACTTTCCCTCGATCTTA 102
43 hTyrThrArgLeuGlnMetGlnGlyGluAlaAlaLeuAlaArgLeu 59
103 CAAAACCTCGACTCCAAATGCAGAGAGAGAGAGAGCTCTGCTCGGTGGCA 152
60 AspGlyAlaArgGluSerAlaProTyrArgGlyMetValArgThrAla 76
153 GAGGTCGCAAGAGAAATCTGCCCTATAGGGAATGTCGACACACCCCT 202
76 uGlyIleIleGluGluGluGlyPheLeuLysLeuTrpGlnGlyValThr 93
203 AGGCATCATTTGAGAGAGAGAGCTTTCTTAAAGCTTTGGCAAGAGTGACAC 252
93 rAlaIleTyrArgHisValValTyrSerGlyGlyArgMetValThrTyr 109
253 CGCCATTTACAGACAGCATGATATTCGTGAGGTGATGTCACATAT 302
110 GluHisLeuArgGluValValPheGlyLysSerGluAsp 122
303 GAACATCTCCGAGAGGTGTGTTGGCAAAAGTGAAGAT 341
seq_name: gb_est78:BE772467

```

```

seq_documentation_block:
LOCUS   BE772467       407 bp      mRNA      EST      20-SEP-2000
DEFINITION   RC2-FT0126-270600-011-d05 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE772467
VERSION     BE772467.1  GI:10226125
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 407)
REFERENCE   1 (bases 1 to 407)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Bionesi,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.B., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
JOURNAL    Contact: Simpson A.J.G.
MEDLINE    Laboratory of Cancer Genetics
COMMENT    Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?fl=st2=RC2-FT0126-270
600-011-d05&ft3=2000-06-27&ft4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 407.
Location/Qualifiers
1..407

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESFES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
```

BASE COUNT 113 a 88 c 83 g 123 t

ORIGIN

alignment_scores:
Quality: 495.00 Length: 95
Ratio: 5.211 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-397-342-1 x BE772467/rev ..

Align seg 1/1 to reverse of: BE772467 from: 1 to: 407

206 G|Y|A|S|P|L|E|U|T|H|T|Y|R|A|S|P|T|H|V|A|L|L|S|T|Y|R|E|U|V|A|L|L|A|S|N|T|H 222
|||||
405 G|G|A|C|A|T|T|A|C|C|A|C|T|T|A|G|A|T|A|C|A|G|T|G|A|A|C|A|C|T|T|G|T|A|T|G|A|A|T|A|C 356
|||||
222 r|P|L|E|U|G|L|U|A|S|P|A|N|L|E|U|T|H|S|G|L|Y|L|E|U|S|E|R|S|E|R|C|Y|S|S|E|R|G 239
|||||
355 A|C|C|A|C|T|G|A|G|G|A|C|A|T|A|T|C|A|G|A|C|T|C|A|G|G|T|T|A|T|C|A|G|T|T|A|T|G|T|C|G 306
|||||
239 l|Y|L|E|U|V|A|L|A|S|E|R|L|E|U|G|L|Y|T|H|R|P|O|A|A|S|P|V|A|L|L|E|Y|S|E|R|A|R|G 255
|||||
305 G|A|C|T|G|T|A|G|C|T|T|A|T|C|T|G|G|A|A|C|A|C|C|G|G|A|T|G|C|A|T|C|A|A|A|A|G|C|A|G 256
|||||
256 l|L|E|U|T|S|N|G|I|N|P|R|O|A|R|G|A|S|P|L|Y|S|G|L|Y|R|G|L|Y|L|E|U|T|Y|R|Y|S|S|E 272
|||||
255 A|T|A|T|G|A|T|A|C|A|C|C|A|G|A|T|A|A|C|A|G|A|G|G|G|A|C|T|T|G|T|A|T|A|A|T|C 206
|||||
272 r|S|E|R|T|H|A|S|P|C|Y|S|L|E|U|I|L|E|G|I|N|A|L|A|G|I|N|G|L|Y|L|E|U|T|Y|R|H|E|T|S|E|R|L 289
|||||
205 A|T|C|A|C|T|G|A|C|G|C|T|T|G|A|T|C|A|G|G|C|T|T|C|A|G|G|C|T|T|C|A|G|A|G|A|T|T|C|A|G|A|C|T|C 156
|||||
289 e|U|T|Y|R|Y|S|G|L|Y|P|H|E|U|P|R|O|S|E|R|T|P|L|E|U|A|R|G|I|Y 300
|||||
155 T|A|T|A|T|A|A|G|C|T|T|T|A|C|A|C|A|C|T|T|G|C|T|G|A|G|A|A|T|G 121
|||||

seq_name: gb_est50:AW729699

seq_documentation_block:
LOCUS AW729699 928 bp mRNA EST 16-NOV-2000
DEFINITION GA_Ea0025M13 Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0025M13, mRNA sequence.
ACCESSION AW729699
VERSION AW729699.1 GI:7627300
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
Magnoliophyta: eudicotyledons, core eudicots, Rosidae, eurosids II;
Malvales: Malvaceae; Gossypium.
1 (bases 1 to 928)
Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
High quality sequence stop: 928.
Location/Qualifiers
1. 928

FEATURES
source
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone_lib="GA_Ea0025M13"
/tissue_type="fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: PBK-CMV; Site: 1: EcoRI; Site: 2: XhoI"

BASE COUNT 223 a 190 c 261 g 254 t

ORIGIN

alignment_scores:
Quality: 492.50 Length: 252
Ratio: 2.706 Gaps: 5
Percent Similarity: 72.222 Percent Identity: 43.254

alignment_block:
US-09-397-342-1 x AW729699 ..

Align seg 1/1 to: AW729699 from: 1 to: 928

73 A|R|G|H|I|A|L|E|U|G|L|Y|L|E|G|L|U|G|L|Y|P|H|E|U|L|Y|S|L|E|U|T|P|G|I 89
|||||
48 C|G|T|G|C|C|C|C|C|G|G|G|A|T|C|G|T|C|G|A|C|C|A|G|C|C|T|C|A|G|G|C|T|G|T|A|A 97
|||||
89 n|G|I|V|A|L|H|R|P|R|O|A|I|L|E|Y|R|A|R|G|H|S|V|A|L|T|Y|R|S|E|R|G|L|Y|R|G|I 106
|||||
98 A|G|T|C|T|C|C|C|C|G|G|G|A|T|T|A|T|T|A|G|C|A|C|T|T|C|T|A|C|A|C|T|C|C|A|T|T|G|C|A 147
|||||
106 e|T|V|A|L|H|T|Y|R|G|L|U|H|S|L|E|U|A|R|G|L|U|V|A|L|P|H|E|G|L|Y|S|E|R|G|L|U|A|S|P 122
|||||
148 T|C|G|C|G|G|T|A|C|G|A|A|T|T|A|G|A|A|T|T|G|T|G|G|.....A|C|G|C|C|G|T 188
|||||
123 G|L|U|H|S|T|Y|R|P|R|O|L|E|U|T|P|L|Y|S|E|R|V|A|L|L|E|G|L|Y|H|E|U|T|A|L|G|I|Y|A 139
|||||
189 G|C|T|C|T|T|T|C|T|G|T|C|T|T|A|A|G|C|G|T|T|A|G|G|T|G|C|A|T|T|C|C|G|G|C|G 238
|||||
139 l|L|E|G|L|Y|P|H|E|U|A|L|A|S|P|R|O|T|H|A|S|P|L|E|U|V|A|L|Y|S|V|A|L|G|I|N|E|T|G 156
|||||
239 T|A|T|G|C|T|A|G|T|T|G|G|C|A|A|T|G|C|A|G|G|C|G|T|T|A|G|T|A|G|T|G|A|G|A|T|C 288
|||||
156 l|M|E|T|G|L|Y|L|Y|S|A|R|G|L|Y|S|L|E|U|G|L|Y|L|Y|S|P|R|O|L|E|U|A|R|G|P|H|E|A|R|G|I|Y 172
|||||
289 A|A|G|G|G|A|T|G|C|G|G|T|T|G|T|A|A|C|A|C|G|G|G|C|T|T|C|A|A|C|C|G|T|C|G|A|T|A|A|A|G|A 338
|||||
173 V|A|H|S|H|S|A|P|H|E|A|L|Y|S|I|L|E|U|A|L|G|I|N|G|L|Y|I|L|E|A|R|G|I|Y|L 189
|||||
339 C|C|G|T|G|A|T|G|C|T|T|T|A|C|A|A|A|T|T|G|A|C|C|A|T|G|G|A|G|G|C|T|T|G|G|A|G|G|T|C 388
|||||
189 e|U|T|P|A|L|A|G|I|Y|R|P|V|A|P|R|O|A|S|N|L|E|G|I|N|A|R|G|A|L|A|L|A|L|E|U|V|A|L|S|N|E|T 205
|||||
389 T|T|T|G|G|A|A|G|G|G|T|T|T|G|C|A|A|T|G|T|G|C|A|G|G|G|C|T|T|A|G|T|A|C|A|C|T|G 438
|||||
206 G|L|Y|A|S|P|L|E|U|T|H|T|Y|R|A|S|P|T|H|V|A|L|L|S|T|Y|R|E|U|V|A|L|L|A|S|N|T|H 222
|||||
439 G|G|T|G|A|C|T|A|G|C|C|T|T|A|G|A|T|C|A|G|C|A|A|A|G|C|T|T|T|G|A|T|A|A|T|A|T|A|C|A 488
|||||
222 r|P|R|O|L|E|U|G|L|U|A|S|P|A|N|L|E|U|T|H|H|S|G|L|Y|L|E|U|S|E|R|S|E|R|C|Y|S|S|E|R|G 239
|||||
489 G|A|T|A|T|G|A|T|A|T|A|C|A|T|T|A|G|C|A|C|A|C|A|C|A|T|T|G|C|A|T|C|A|T|G|T|G|A|G 538
|||||
239 l|Y|L|E|U|V|A|L|A|S|E|R|L|E|U|G|L|Y|T|H|R|P|O|A|A|S|P|V|A|L|L|E|Y|S|E|R|A|R|G 255
|||||
539 G|T|C|T|C|T|C|T|A|C|A|A|T|G|T|G|T|G|C|G|G|C|T|G|A|T|T|G|T|G|T|A|A|A|C|A|A|G 588

High quality sequence stop: 393.
Location/Qualifiers
1. 393
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate; tumor; Vector: puc18; Site: 1; SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 109 a 85 c 79 g 120 t
ORIGIN

alignment_scores:
Quality: 485.00 Length: 95
Ratio: 5.160 Gaps: 0
Percent Similarity: 98.947 Percent Identity: 98.947

alignment_block:
US-09-397-342-1 x BE772458/rev ..
Align seg 1/1 to reverse of: BE772458 from: 1 to: 393

206 GlysplethrThrTyrAspThrValylshsTYrLeuValLeuAsnTh 222
|||||
391 GGAGATTAAACCACTTAACTGAAACACTGTTGTTGAATAC 342
222 rProleugluaspasnilemetThHisGlyleuSerleucysSerg 239
|||||
341 ACCACTTGAAGCAATATCATGATCAGCTGTTATCAAGTTATATGTTCTG 292
239 lyleuValAlaSerlleuGlyThrProAlaaspValilleysSerArg 255
|||||
291 GACTGTGAGCTTCTATTTCTGGGAACACCGCATGTCATCAAAAGCAGA 242
256 llemetasnglnProArgaspLysglnGlyArgGlyLeuLeuTyrLysSe 272
|||||
241 ATAATGAATCAACTACGAGATTAACAAGGAGGACTTTGTATTAATC 192
272 rSerThrAspCysleuIleGlnAlaValGlnGlyGlnGlyPheMetSerL 289
|||||
191 ATCGACTGACTGCTTGAATTCAGGCTGTTCAAGGTAAAGATTCAATGATC 142
289 eutyrlYsglyPheleupProSerTPrleuArgMet 300
|||||
141 TATATAAAGCGTTTACCATCTTGCGTGAAGATG 107

seq_name: qb_esc78:BE772464

seq_documentation_block:
LOCUS BE772464 412 bp mRNA EST 20-SEP-2000
DEFINITION RC2-FT0126-270600-011-c01 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772464
VERSION BE772464.1 GI:10226122
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunsteht,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
/M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=6t2-RC2-FT0126-270
600-011-c01&t3=2000-06-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 28
High quality sequence stop: 412.
Location/Qualifiers
1..412
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate; tumor; Vector: puc18; Site: 1; SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 113 a 90 c 86 g 123 t
ORIGIN

alignment_scores:
Quality: 481.00 Length: 130
Ratio: 4.859 Gaps: 1
Percent Similarity: 76.154 Percent Identity: 74.615

alignment_block:
US-09-397-342-1 x BE772464/rev ..
Align seg 1/1 to reverse of: BE772464 from: 1 to: 412

206 GlysplethrThrTyrAspThrValylshsTYrLeuValLeuAsnTh 222
|||||
410 GGAGATTAAACCACTTAACTGAAACACTGTTGTTGAATAC 361
222 rProleugluaspasnilemetThHisGlyleuSerleucysSerg 239
|||||
360 ACCACTTGAAGCAATATCATGATCAGCTGTTATCAAGTTATATGTTCTG 311
239 lyleuValAlaSerlleuGlyThrProAlaaspValilleysSerArg 255
|||||
310 GACTGTGAGCTTCTATTTCTGGGAACACCGCATGTCACCAAAAGCAGA 261
256 llemetasnglnProArgaspLysglnGlyArgGlyLeuLeuTyrLysSe 272
|||||
260 ATAATGAATCAACTACGAGATTAACAAGGAGGACTTTGTATTAATC 211
272 rSerThrAspCysleuIleGlnAlaValGlnGlyGlnGlyPheMetSerL 289
|||||
210 ATCGACTGACTGCTTGAATTCAGGCTGTTCAAGGTAAAGATTCAATGATC 161
289 eutyrlYsglyPheleupProSerTPrleuArgMet..... 300
|||||
160 TATATAAAGCGTTTACCATCTTGCGTGAAGATGTATGTCGATC 111
300 300
110 TCACAGGCATCCAAATGAGACCCCGACAGCATTTTCAAGAAGAATGCA 61
301ThpTrpSerMetValPheTrpLeu 309

60 AGCCTGACCACTTTCACCTTGGGCGAGAGAGTTTGGCCTT 21

|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 16:10:21 ; Search time 1145.32 seconds
(without alignments)
7921.668 Million cell updates/sec

Title: US-09-397-342-2
1039
Sequence: 1 ccgagccgcgagcccgctat.....cagatatccatcacactgc 1039

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues
Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:.*
2: gb_est2:.*
3: gb_est3:.*
4: gb_est4:.*
5: gb_est5:.*
6: gb_est6:.*
7: gb_est7:.*
8: gb_est8:.*
9: gb_est9:.*
10: gb_est10:.*
11: gb_est11:.*
12: gb_est12:.*
13: gb_est13:.*
14: gb_est14:.*
15: gb_est15:.*
16: gb_est16:.*
17: gb_est17:.*
18: gb_est18:.*
19: gb_est19:.*
20: gb_est20:.*
21: gb_est21:.*
22: gb_est22:.*
23: gb_est23:.*
24: gb_est24:.*
25: gb_est25:.*
26: gb_est26:.*
27: gb_est27:.*
28: gb_est28:.*
29: gb_est29:.*
30: gb_est30:.*
31: gb_est31:.*
32: gb_est32:.*
33: gb_est33:.*
34: gb_est34:.*
35: gb_est35:.*
36: gb_est36:.*
37: gb_est37:.*
38: gb_est38:.*
39: gb_est39:.*
40: gb_est40:.*
41: gb_est41:.*
42: gb_est42:.*
43: gb_est43:.*
44: gb_est44:.*
45: gb_est45:.*
46: gb_est46:.*
47: gb_est47:.*
48: gb_est48:.*
49: gb_est49:.*
50: gb_est50:.*
51: gb_est51:.*
52: gb_est52:.*
53: gb_est53:.*
54: gb_est54:.*
55: gb_est55:.*
56: gb_est56:.*
57: gb_est57:.*
58: gb_est58:.*
59: gb_est59:.*
60: gb_est60:.*
61: gb_est61:.*
62: gb_est62:.*
63: gb_est63:.*
64: gb_est64:.*
65: gb_est65:.*
66: gb_est66:.*
67: gb_est67:.*
68: gb_est68:.*
69: gb_est69:.*
70: gb_est70:.*
71: gb_est71:.*
72: gb_est72:.*
73: gb_est73:.*
74: gb_est74:.*
75: gb_est75:.*
76: gb_est76:.*
77: gb_est77:.*
78: gb_est78:.*
79: gb_est79:.*
80: gb_est80:.*
81: gb_est81:.*
82: gb_est82:.*
83: gb_est83:.*
84: gb_est84:.*
85: gb_est85:.*
86: gb_est86:.*
87: gb_est87:.*
88: gb_est88:.*
89: gb_est89:.*
90: gb_est90:.*
91: gb_est91:.*
92: gb_est92:.*
93: gb_est93:.*
94: gb_est94:.*
95: gb_est95:.*
96: gb_est96:.*
97: gb_est97:.*
98: gb_est98:.*
99: gb_est99:.*
100: gb_est100:.*
101: gb_est101:.*
102: gb_est102:.*
103: gb_est103:.*
104: gb_est104:.*
105: gb_est105:.*
106: gb_est106:.*
107: gb_est107:.*
108: gb_est108:.*
109: gb_est109:.*
110: gb_est110:.*
111: gb_est111:.*
112: gb_est112:.*
113: gb_est113:.*
114: gb_est114:.*
115: gb_est115:.*
116: gb_est116:.*

44: em_esthum10:.*
45: em_esthum11:.*
46: em_esthum12:.*
47: em_esthum13:.*
48: em_esthum14:.*
49: em_esthum15:.*
50: em_esthum16:.*
51: em_esthum17:.*
52: em_esthum18:.*
53: em_esthum19:.*
54: em_esthum20:.*
55: em_esthum21:.*
56: em_esthum22:.*
57: em_esthum23:.*
58: em_esthum24:.*
59: em_esthum25:.*
60: em_esthum26:.*
61: em_esthum27:.*
62: em_esthum28:.*
63: em_esthum29:.*
64: em_esthum30:.*
65: em_esthum31:.*
66: em_esthum32:.*
67: em_esthum33:.*
68: em_esthum34:.*
69: em_esthum35:.*
70: em_esthum36:.*
71: em_esthum37:.*
72: em_esthum38:.*
73: em_esthum39:.*
74: em_esthum40:.*
75: em_esthum41:.*
76: em_esthum42:.*
77: em_esthum43:.*
78: em_esthum44:.*
79: em_esthum45:.*
80: em_esthum46:.*
81: em_esthum47:.*
82: em_esthum48:.*
83: em_esthum49:.*
84: em_esthum50:.*
85: em_esthum51:.*
86: em_esthum52:.*
87: em_esthum53:.*
88: em_esthum54:.*
89: em_esthum55:.*
90: em_esthum56:.*
91: em_esthum57:.*
92: em_esthum58:.*
93: em_esthum59:.*
94: em_esthum60:.*
95: em_esthum61:.*
96: em_esthum62:.*
97: em_esthum63:.*
98: em_esthum64:.*
99: em_esthum65:.*
100: em_esthum66:.*
101: em_esthum67:.*
102: em_esthum68:.*
103: em_esthum69:.*
104: em_esthum70:.*
105: em_esthum71:.*
106: em_esthum72:.*
107: em_esthum73:.*
108: em_esthum74:.*
109: em_esthum75:.*
110: em_esthum76:.*
111: em_esthum77:.*
112: em_esthum78:.*
113: em_esthum79:.*
114: em_esthum80:.*
115: em_esthum81:.*
116: em_esthum82:.*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: em_gss_pln1:*
 191: em_gss_pln2:*
 192: em_gss_pro:*
 193: em_gss_rdd1:*
 194: em_gss_rdd2:*
 195: em_gss_rdd3:*
 196: em_gss_rdd4:*
 197: em_gss_rdd5:*
 198: em_gss_vrt1:*
 199: em_gss_vrt2:*
 200: em_gss_vrt3:*
 201: gb_gss1:*
 202: gb_gss2:*
 203: gb_gss3:*
 204: gb_gss4:*
 205: gb_gss5:*
 206: gb_gss6:*
 207: gb_gss7:*
 208: gb_gss8:*
 209: gb_gss9:*
 210: gb_gss10:*
 211: gb_gss11:*
 212: gb_gss12:*
 213: gb_gss13:*
 214: gb_gss14:*
 215: gb_gss15:*
 216: gb_gss16:*
 217: gb_gss17:*
 218: gb_gss18:*
 219: gb_gss19:*
 220: gb_gss20:*
 221: gb_gss21:*
 222: gb_gss22:*
 223: gb_gss23:*
 224: gb_gss24:*
 225: gb_gss25:*
 226: gb_gss26:*
 227: gb_gss27:*
 228: gb_gss28:*
 229: gb_gss29:*
 230: gb_gss30:*
 231: gb_gss31:*
 232: gb_gss32:*
 233: gb_gss33:*
 234: gb_gss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	508.2	48.9	584	111	AW162249
C 2	482	46.4	483	111	AW157357
C 3	482	46.4	483	111	AW157496
C 4	477.8	46.0	633	111	AW160749
C 5	427	41.1	654	31	AV652060
C 6	416.2	40.1	564	111	AW161176
C 7	369.2	35.5	526	103	AI878851
C 8	334.2	32.2	341	152	F11430
C 9	287	27.6	407	139	BE772467
C 10	284.4	27.4	393	139	BE772458
C 11	283.8	27.3	412	139	BE772464
C 12	282.8	27.2	331	139	BE773393
C 13	272.4	26.2	409	139	BE772459
C 14	269.4	25.9	391	139	BE772466
C 15	243.6	23.4	607	162	BE061032
C 16	239.8	23.1	803	111	AW147976
C 17	237.8	22.9	590	171	BF954659
C 18	225.6	21.7	375	111	AW161833
					AW162249 au75611.x
					AW157357 au88a06.x
					AW157496 au79409.x
					AW160749 au75611.y
					AV652060 AV652060
					AW161176 au79409.y
					AI878851 au50602.y
					F11430 HSC2WH061.n
					BE772459 RC2-FT012
					BE772467 RC2-FT012
					BE772458 RC2-FT012
					BE772464 RC2-FT012
					BE773393 QV1-FT017
					BE772459 RC2-FT012
					BE772466 RC2-FT012
					BE061032 QV0-BT004
					AW147976 da01008.x
					BF954659 MR3-NN021
					AW161833 au88a06.y

[illegible]

ALIGNMENTS

RESULT	1
AM162249/c	
LOCUS	
DEFINITION	
AM162249	584 bp mRNA
au75ell.x1 schneider fetal brain 00004 Homo sapiens cDNA clone	EST
IMAGE:278B224.3, similar to SW:UCP4_HUMAN O95847 MITOCHONDRIAL	09-NOV-1999
UNCOUPLING PROTEIN 4 ; , mRNA sequence.	

ACCESSION	AM162249	
VERSION	AM162249.1	GI:6301282
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE 1 (bases 1 to 584)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE	JOURNAL	COMMENT
WashU-NCI human EST pr	Unpublished (1997)	other_ESTS: au75e11.y1

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.lnl.gov) for further information
Seq primer: -40UP from Glbco
High quality sequence stop: 463.
Location/Qualifiers

FEATURES	
source	Location/Qualifiers
1..584	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:2782124"	
/clone_1ib="Schneider fetal brain 00004"	
/sex="male"	
/tissue_type="frontal lobe"	
/dev_stage="5 months post-conception"	
/lab_host="DH10B"	

/Note="Organ: brain; Vector: pluescript SK (Stratagene);
Site_1: SclI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGGAGGAGAAGAGCTCAGAGATCCTTAATTAAATTAAATTC
CCCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGGAGGAGAGACTCGAGTGTGTGTGTGTGTGTGT-3'.
The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."
121 a 182 c 138 g 143 t

Query Match	48.9%;	Score 508.2;	DB 11;	Length 584;
Best Local Similarity	99.4%;	Pred. No. 1.6e-141;		
Matches 510;	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps	0

Q7 9 ggaatcccggtatcgccttctgagcgtactgcctgaatatgccgacccggagagagaagacct 68
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
D8 514 GGAGGCGCTATTCGTCTTCCGGCTACTACGCAGATATCCTCCGCCGAGAGAGAGAGAGGCT 455

QY 69 ttgccgtgaccagagatggcccgagcgagcaattcctactgtccgctgcgcgc 128

[illegible]

xy lz cacccgugcagacaaaccclcccccguacllcccaaaaacccgaclcccaadlcga lbb

Db 394 TACCGTGCCGAGCTAGCAACCTTTCCTGGATCTCACAAAACTGACTCCAATGCA 335

QY 189 aqqaqaacacactcttqactcqqtlqqaaacacqqtqcaaaqaatatctqccccctataqqq 248

[illegible]

249 aatggtgcgcacagccctagggaatcattgaagaggaaagccttcttaagccttggcaagg 308

Db 274 AATGGTGCACAGCCCTAGGATCATTTGAAGAGGAAGCTTCTAAAGCTTGGCAAG 215

309 atgacacccaccatttacagacacgtatgtatatcttgaagtcgaatggtcacatatga 368

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

369 acatctccgagaggttggttgcaaaagtgaagatgagcattatccctltggaatc 428

Db 154 ACATCTCCGAGAGTTGTGTTGGCAAAAGTGAAGATGAGCATTAATCCCTTGAATC 95

Ov A29 atcattacaaatgaatactaatattatataccacatttttaaccacatccaaactaacnt 488

[illegible]

QY 489 agtgaaggttcagatgcaaatggaaggaag 521

Db 34 AGTGAAGGTCAGATGCAATGGAAGGAAAAG 2

RESULT 2

AW157357	483 bp	MPNA	EST	04-NOV-1999
LOCUS				
AW157357/C				

DEFINITION au88a06.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone

UNCOUPLING PROTEIN 4 ; mRNA sequence.

ACCESSION AM157357
 VERSION AM157357 1 CT-6000250

KEYWORDS EST.

SOURCE *humana*.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 483)
Mammalia; Eulnidae; Primates; Catarrhini; Hominoidea; Homo.

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 654)
AUTHORS Qian,B., Wu,T., Huang,Q., Kang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens CDNA clone
JOURNAL Unpublished (2000)
COMMENT Contact: Zengang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCCWD06"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 213 a 102 c 163 g 171 t 5 others
ORIGIN

Query Match 41.1%; Score 427; DB 31; Length 654;
Best Local Similarity 100.0%; Pred. No. 4e-117;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 514 ggaagaaagaaactggaagaaacacatgcatctgctggtgacatcatgcatctgca 573
DB 1 GGAAAAAGAACTGGAAGAAACACATTCGATTCGTGGTACATCATGATTTGCA 60
QY 574 aaatctgctggaaggaatcaagagcttggcgagctggtggtaccatataca 633
DB 61 AAAATCTTAGCTGAAGAGAAATACGAGGCTTTGGGCGAGCTGGGTACCAATATACAA 120
QY 634 aagaagacagctggaatggaatggaatggaatggaatggaatggaatggaatgga 693
DB 121 AAGAGCAGCACTGGTGAATATGGAGATTATACCACTTATGATACAGTAAACACTACTTG 180
QY 694 gtaatgatacaccacttgaagacaataatcatgactgacgttcaatgattatgtct 753
DB 181 GATTGAATATACCACTTGAAGACATATCATGACACAGGTTTATCAAGTTATCTTCT 240
QY 754 ggaactgtagcttctattcttgggaacacagccgagatgcatcaaaagcaaatgaa 813
DB 241 GCAGCTGGTGGTCTATCTTCTGGGAACACCGAGTCGATCAAAAAGCAGAAATGAAT 300
QY 814 caacacagagataaacaagaaggagacttggtaataatcagctgactgactgatt 873
DB 301 CAACCAAGGATTAACAAAGAGGAGACTTTGTATTAATATCACTGACTGCTGATT 360
QY 874 cagctgtctcaagtgaaagatcatgactatataaaagcttctacacatctgagct 933
DB 361 CAGGCTGTCAAGGTGAAGATTCAATGATATATAAAGCTTTTACCATCTTGCTG 420
QY 934 agaatga 940
DB 421 AGAATGA 427

RESULT 6

AW161176 564 bp mRNA EST 09-NOV-1999
LOCUS
DEFINITION au79d09.y1 schneider fetal brain 00004 Homo sapiens CDNA clone
IMAGE:2782481.5' similar to SW:U0P4.HUMAN 095847 MITOCHONDRIAL
UNCOUPLING PROTEIN 4 ; mRNA sequence.
ACCESSION AW161176 GI:6300209
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 364)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Maira,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: au79d09.x1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
High quality sequence stop: 435.

FEATURES
source
1..564
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2782481"
/clone_lib="Schneider fetal brain 00004"
/sex="Male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene); Site_1: SbfI; Site_2: XhoI; Double-stranded cDNA was prepared from human fetal brain tissue. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence:
5'-GAGAGAGAGAGAGCTCAAGATCCTTAATTAATTAATCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGCTCAAGTCTTTTCTTTTCTTTT-3'. The library was size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy)."
BASE COUNT 129 a 148 c 180 g 107 t
ORIGIN

Query Match 40.1%; Score 416.2; DB 111; Length 564;
Best Local Similarity 99.3%; Pred. No. 6.7e-114;
Matches 418; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 gataccggttactgcttgcctgactgtaatgctcgtcccgagagagagagagct 68
DB 144 GGAGTGGCTTATCGTTCCTGCTACTGCTGATGTCGTCGCCGAGAGAGAGAGGCT 203
QY 69 ttggcgtgacccagagatgagcccgagagagaaatctcactgctcgcgtgcgcgc 128
DB 204 TTTGGCGCTGAGCCAGAGATGGCCCCGAGCGAATTCCTACCTGCTGCGCGCGC 263
QY 129 taccggtgcccagctagcaaccttccctgtagtctcaaaaactgactcgaatgca 188
DB 264 TACCGTGGCGAGCTAGCACTTTCCTCGATCTCACAAAACCTCGACTCCAATGCA 323

189 aggaagaacagctctgtcgtggtggaagagtgcaagaagaatctgccccctataagg 248
 |||||||
 Db 324 AGGAGAACAGACTCTTGGCTTGGGAGACGCTGCAAGAGAAATCTGCCCTATAGGG 383
 249 aatggtgcacacagccctagagatcatgaagaagaagcttctcaagcttggcaag 308
 Db 384 AATGTTGGCGACAGCCCTAGGGATCATTTGAAGAGAGAGCGTTTAAAGCTTGGCAAG 443
 309 agtgcaccccgccattacagacagctagtattcttggaggtgcaatggtacatatga 368
 Db 444 AGTGACACCCCGCATTTACAGACAGCTAGTATTCTGGAGCTCGAATGCTCACAATGGA 503
 369 acatctccgaagagtggtgttggcaaaagtgaaagatgacatctcccttggaaatc 428
 Db 504 ACATCTCCGAGAGTGTGTGTTGGCAAAAGTGAAGATGAGCATTTCCCTTTGGAATC 563
 429 a 429
 564 A 564

RESULT 7
 A1878851 526 bp mRNA EST 23-AUG-1999
 LOCUS aus0e02.y1 schneider fetal brain 00004 Homo sapiens cDNA clone
 DEFINITION IMAGE:2518202.5', similar to WP:K07B1.3 CELL1880 PROTON-CHANNEL
 PROTEIN,, mRNA sequence.
 A1878851
 VERSION A1878851.1 GI:5552900
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 526)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,
 J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B.,
 White,X., Wylie,T., Waterston,K. and Wilson,R.
 WashU-NCI human EST Project
 Unpublished (1997)
 CONTACT: Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40RP from G1bco
 High quality sequence stop: 417.
 Location/Qualifiers
 1..526
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2518202"
 /clone_1bp="Schneider fetal brain 00004"
 /sex="male"
 /tissue_type="frontal lobe"
 /dev_stage="5 months post-conception"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript SK (Stratagene);
 Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
 prepared from human fetal brain tissue. 5' and 3'
 adaptors were used in cloning as follows: 5' adaptor
 sequence:
 5'-GAGAGAGAGAGAGCTCAGAGATCCTTAATTAATTAATATCCCCCCCCCCC-3'
 and 3' adaptor sequence:
 5'-GAGAGAGAGAGCTCAGAGTCTTTTCTTTTCTTTTCTTTT-3'. The library was
 size-selected for >0.5 kb inserts and has an average
 insert size estimated at 1.2 kb. This library was
 constructed using the CAP-trapper method for full-length

enrichment and has not undergone amplification. Library
 was constructed by Dr. Claudio Schneider (LNCIB-Area
 Science Park, Trieste, Italy).
 BASE COUNT 115 a 141 c 173 g 96 t 1 others
 ORIGIN

Query Match 35.5%; Score 369.2; DB 103; Length 526;
 Best Local Similarity 98.9%; Pred. No. 8.6e-100;
 Matches 371; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

9 gataccgctatcgtcttgcgtctgtaatgctccgccgagaggaagaagct 68
 |||
 Db 144 GGAGTGCCTTATTCGCTTGGCTGCTGCTGATGTCCTCCGAGAGAGAGAGAGCT 203
 69 ttgcgcgtgacacagagatgagcccgagcgagaaatctactctgcgtgcgcgc 128
 Db 204 TTTGCCGCTGACCCAGAGATGGCCCCGAGCGAGCAAAATTCCTACTGTCGGCTCGCGCG 263
 129 taccgtggccgagctagacaaccttccctggtatctcaaaaactgactccaatga 188
 |||||||
 Db 264 TACGCTGGCCGAGCTAGACACCTTCCCTGGATCTCAAAAACCTCACTCAATATCA 323-
 189 aggaagaacagctctgtcgtggtggaagagtgcaagaagaatctgccccctataagg 248
 |||||||
 Db 324 AGGAGAACAGACTCTTGGCTTGGGAGACGCTGCAAGAGAAATCTGCCCTATAGGG 383
 249 aatggtgcacacagccctagagatcatgaagaagaagcttctcaagcttggcaag 308
 Db 384 AATGTTGGCGACAGCCCTAGGGATCATTTGAAGAGAGAGCGTTTAAAGCTTGGCAAG 443
 309 agtgcaccccgccattacagacagctagtattcttggaggtgcaatggtacatatga 368
 |||||||
 Db 444 AGTGACACCCCGCATTTACAGACAGCTAGTATTCTGGAGCTCGAATGCTCACAATGGA 503
 369 acatctccgaagagtggtgttggcaaaagtgaaagatgacatctcccttggaaatc 428
 |||||||
 Db 504 ACATCTCCGAGAGTGTGTGTTGGCAAAAGTGAAGATGAGCATTTCCCTTTGGAATC 563

RESULT 8
 F11430 341 bp mRNA EST 12-MAR-1995
 LOCUS HSC2WH061 normalized infant brain cDNA Homo sapiens cDNA clone
 DEFINITION C-2wh06, mRNA sequence.
 F11430
 VERSION F11430.1 GI:705728
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 341)
 Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,
 'M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
 Sebastiani,Kadaktchis,C. and Tessier,A.
 IMAGE: molecular integration of the analysis of the human genome
 and its expression
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 95277534
 Contact: Genethon
 Genexpress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read.
 Genexpress_library_id: C; Genexpress_sequence_id: y1c-2wh06
 Seq primer: (-21)M13 universal.
 Location/Qualifiers
 1..341

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2zh06"
/clone_lib="normalized infant brain cDNA"
/sex="female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: latmid BA; Site: 1: HindIII;
Site: 2: NotI; sex: female; dev: stage=3 months old;
isolate=muscular atrophy patient; tissue type=total brain
; total mRNA was oligo (dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
latmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization method:
Bento Soares, P.N.A.S in press"

BASE COUNT      86 a      85 c      96 g      72 t      2 others
ORIGIN

Query Match      32.2%; Score 334.2; DB 152; Length 341;
Best Local Similarity 98.5%; Pred. No. 2.4e-89;
Matches 336; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 65 gacttttcgctgacccgagatggtcccgagcgagcaaatctcactgtccgctgag 124
|||||
DB 1 ggcctttggccgctgacccgagatggtcccgagcgagcaaatctcactgtccgctgag 60
|||||

QY 125 cggctacgctggtccgagctgagcaacgttcctcgtgattctcacaanaactcgactcaca 184
|||||
DB 61 cggctacgctggtccgagctgagcaacgttcctcgtgattctcacaanaactcgactcaca 120
|||||

QY 185 tgcagaagaaagcaagctctctgctcgtgttggaagcgttgcaagaagatctgccccctata 244
|||||
DB 121 tgcagaagaaagcaagctctctgctcgtgttggaagcgttgcaagaagatctgccccctata 180
|||||

QY 245 ggggaatggtgagcaagcctcctagagatcattgaagaagaaagccttctcaaaagcttgagc 304
|||||
DB 181 ggggaatggtgagcaagcctcctagagatcattgaagaagaaagccttctcaaaagcttgagc 240
|||||

QY 305 aaggaatgcaaccccgcatcttaccagacacgtatgtaattctggaagtcgaatgacat 364
|||||
DB 241 aaggaatgcaaccccgcatcttaccagacacgtatgtaattctggaagtcgaatgacat 300
|||||

QY 365 atgaacatctccgagaggtgtgtgttggaagaagtgaaagt 405
|||||
DB 301 atgaacatctccgagaggtgtgtgttggaagaagtgaaagt 341
|||||

RESULT 9
LOCUS 772467/c
DEFINITION RC2-FT0126-270600-011-d05 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772467
VERSION BE772467.1 GI:10226125
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 407)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

```

```

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=kt2-RC2-FT0126-270
600-011-d05&it3=2000-06-27&it4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 407.
Location/Qualifiers
1. 407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2zh06"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate tumor; Vector: puc18; Site: 1: SmaI;
Site: 2: SmaI; A mini-library was made by cloning products
derived from ORESSES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      113 a      88 c      83 g      123 t
ORIGIN

Query Match      27.6%; Score 287; DB 139; Length 407;
Best Local Similarity 100.0%; Pred. No. 3.9e-75;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 653 tgggaatttaaccacttctgatacagtgaaacacttggtgattgaatacaccactg 712
|||||
DB 407 tgggaatttaaccacttctgatacagtgaaacacttggtgattgaatacaccactg 348
|||||

QY 713 aggaacaatcctgactcaggttttcaagttatgtctgactgtgactgttctatc 772
|||||
DB 347 aggaacaatcctgactcaggttttcaagttatgtctgactgtgactgttctatc 288
|||||

QY 773 tgggaacccagcgcgtgtgcatcaaaagcagaataatgatacacaagagataaacaag 832
|||||
DB 287 tgggaacccagcgcgtgtgcatcaaaagcagaataatgatacacaagagataaacaag 228
|||||

QY 833 gaaggagactttgtataatcatcagcagctgactgtgattcaaggtgtcaagtgaaag 892
|||||
DB 227 gaaggagactttgtataatcatcagcagctgactgtgattcaaggtgtcaagtgaaag 168
|||||

QY 893 gattcatgagctcatalaaagccttttaccacatctgtgctgagaatg 939
|||||
DB 167 gattcatgagctcatalaaagccttttaccacatctgtgctgagaatg 121
|||||

RESULT 10
LOCUS BE772458 393 bp mRNA EST 20-SEP-2000
DEFINITION RC2-FT0126-270600-011-a03 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772458
VERSION BE772458.1 GI:10226116
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 393)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

```


JOURNAL
MEDLINE
COMMENT

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RC2-FT0126-270
600-011-a036t3-2000-06-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 393.
Location/Qualifiers
1. 393

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate; tumor; Vector: puc18; Site: 1: Sma1;
Site: 2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 109 a 85 c 79 g 120 t
ORIGIN

Query Match 27.4%; Score 284.4; DB 139; Length 393;
Best Local Similarity 99.7%; Pred. No. 2.3e-74;
Matches 285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 654 gggaagatttaacacattatagatagatgaacacactctgtgtatgaatacaccacttga 713
DB 392 GGGGAGATTAAACCACTTATATACAGTAACACACTACTTGTATGTATGATACACACTTGA 333
QY 714 ggacaataatagatcagcaggttatacaagttatgtctcgagctgtgactctatct 773
DB 332 GGACAAATATATAGCTCAGCGTTTATCAAGTTTATGTTCTGGAGCTGATGCTTATCT 273
QY 774 gggaacacacagcagatgtcatcaaaagcagaataatgaatcaacacacagagaataacaag 833
DB 272 GGGAAACACAGCGAGTGTATCAAAAGCAGAAATGAATCAACTACAGATTAACAAG 213
QY 834 aaagggaactttgtataatcatcagactgactgtgattcgaagctgttcaaggtgaag 893
DB 212 AAGGGGACTTTTGTATTAATCATCGACTGACTGCTTATTCAGGCTGTCAAGGTGAAG 153
QY 894 attcatagactctataaagacttttacatcttgctgaagatg 939
DB 152 ATTCAATGAGTCTATATTAAGGCTTTTACCACTTGGCTGGAATG 107

RESULT 11
LOCUS BE772464 412 bp mRNA EST 20-SEP-2000
DEFINITION RC2-FT0126-270600-011-c01 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772464
VERSION BE772464.1 GI:10226122
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 412)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,

TITLE
JOURNAL
MEDLINE
COMMENT

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RC2-FT0126-270
600-011-c01&t3-2000-06-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 28
High quality sequence stop: 412.
Location/Qualifiers
1. 412

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate; tumor; Vector: puc18; Site: 1: Sma1;
Site: 2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 113 a 90 c 86 g 123 t
ORIGIN

Query Match 27.3%; Score 283.8; DB 139; Length 412;
Best Local Similarity 99.3%; Pred. No. 3.5e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 653 ttggaagatttaacacattatagatagatgaacacactctgtgtatgaatacaccactt 712
DB 412 TGGGAGATTAAACCACTTATATACAGTAACACACTACTTGTATGTATGATACACACTTG 353
QY 713 aggaacaataatagatcagcaggttatacaagttatgtctcgagctgtgactctatct 772
DB 352 AGGCAATATATAGCTCAGCGTTTATCAAGTTTATGTTCTGGAGCTGATGCTTATCT 293
QY 773 ttggaacacacagcagatgtcatcaaaagcagaataatgaatcaacacacagagaataacaag 832
DB 292 TGGGAACACAGCGAGTGTATCAAAAGCAGAAATGAATCAACTACAGATTAACAAG 233
QY 833 gaagggaactttgtataatcatcagactgactgtgattcgaagctgttcaaggtgaag 892
DB 232 GAAGGGGACTTTTGTATTAATCATCGACTGACTGCTTATTCAGGCTGTCAAGGTGAAG 173
QY 893 gattcatagactctataaagacttttacatcttgctgaagatg 939
DB 172 GATTCAATGAGTCTATATTAAGGCTTTTACCACTTGGCTGGAATG 126

RESULT 12
LOCUS BE773993 331 bp mRNA EST 20-SEP-2000
DEFINITION QY1-FT0170-100700-268-f04 FT0170 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE773993
VERSION BE773993.1 GI:10227651
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2-qv1-FT0170-100700-268-f04&ts=2000-07-10&tl=1)
Seq primer: puc 18 forward
High quality sequence stop: 331.
Location/Qualifiers

FEATURES
Source
1. 331
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0170"
/dev_stage="Adult"
/note="Organ: prostate_tumor; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 94 a 87 c 52 g 98 t
ORIGIN

Query Match 27.2%; Score 282.8; DB 139; Length 331;
Best Local Similarity 99.3%; Pred. No. 6,6e-74;
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

332 acgtatgtatcttctggaagtcgaatggtacataatgaacatctcgaagagttgttg 391
11 |||||
290 ACCCAGTGTATCTCGAGGTGATGTCACATATGACATCTCGAGAGTGTGTTTG 231
0Y 332 gcaaaaggaagatgagcatatcccccttggaaatcaagtcattggaggatgatgctg 451
Db 230 GCAAAAGTCMAATGACGCTTATCCCTTGGAAATCACTCATTTGGAGCGATGCGCTG 171
0Y 452 ggttatgagcaagtttaagccaatccaactgacctagtgaaagtttcagatgcaaatg 511
Db 170 GGTATTATGGCCAGTTTATAGCCAAATCCAGTGAACCTTAAGTGAAGGTTCAGATGCAAAATG 111
0Y 512 aaggaagaaagaaactggaagaaacacatctgcatcttctgtgtgtacatctgcatctg 571
Db 110 AAGGAAAAAGAACTGGAAGGAAAAACATTCGATTTGCTGTGTACATCATGCAATTTG 51
0Y 572 caaaaacttaagctgaaggaagaatacgaagagcttggcgaagctg 617
11 |||||
Db 50 CAAAATCTTACCTGAAGGAGGAATACGAGGCTTTGGCAGGCTG 5

LOCUS BE772459 409 bp mRNA EST 20-SEP-2000
DEFINITION RC2-FT0126-270600-011-a07 FT0126 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE772459
VERSION BE772459.1 GI:10226117
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2-RC2-FT0126-270600-011-a07&ts=2000-06-27&tl=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 409.
Location/Qualifiers

FEATURES
Source
1. 409
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FT0126"
/dev_stage="Adult"
/note="Organ: prostate_tumor; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 126 a 81 c 90 g 112 t
ORIGIN

Query Match 26.2%; Score 272.4; DB 139; Length 409;
Best Local Similarity 97.6%; Pred. No. 9,4e-71;
Matches 287; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

647 tgaataggaagatttaacacttatgatacagtgaaacactact-ggatgaataca 705
11 |||||
Db 1 TGAATATGGCAATTAACACATATATGATAGACGTGAACACACTTATGATGATACA 60
0Y 706 ccaacttgagacaataatacgaactcaaggttatacaagttatgtctgagctgtagct 765
11 |||||
Db 61 CCACCTTGAAGCAAAATATCATGACTCAGCGTTTATCAAGTTTATGTTTCGACTGTAGCT 120
0Y 766 tctattctggagaaacccagcgatgcatcaaaagcagaataatgaatacaaccagagat 825
11 |||||
Db 121 TCTATTCTGGGAACACACACCGATGTCATCAAAACAGATATATCAATCAACCGAGAT 180
0Y 826 aaacaagaagaagagcttctgataaatcatcagatgagctgttgaatcaagctttcaa 885
11 |||||
Db 181 AAACAAGGAAGGCACTTTTGTATTAATCATGACTGACTGCTTGATTCAGGCTGTCAA 240
0Y 886 gttgaagatcatgagctatataaaggcttttaaccatcttgctgagagatg 939

Db 241 GGTAAGCATCATGACTATATTAAGGCTTTTACCACTTGGCTGAGATG 294

|||||

RESULT 14

LOCUS BE772466 391 bp mRNA EST 20-SEP-2000

DEFINITION RC2-PT0126-270600-011-d04 PT0126 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE772466

VERSION BE772466.1 GI:10226124

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 391)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.U.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC2-PT0126-270600-011-d04&ts=2000-06-27&td=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 391.
Location/Qualifiers
1. 391
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0126"
/dev_stage="Adult"
/note="Organ: prostate; tumor; Vector: puc18; Site: 1; SmaI; Site: 2; SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 118 a 80 c 87 g 106 t

ORIGIN

Query Match 25.9%; Score 269.4; DB 139; Length 391;
Best local Similarity 99.6%; Pred. No. 7.4e-70;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 669 ttatgatacgtgaacactgtgtatgataacacactgagagacaataatcagac 728

|||||

Db 6 TTATGATACAGTGAACACACTAGGATGATGAATACACACTGAGGACATATCATCAGAC 65

|||||

QY 729 tcaagctttatcaagttatgtctcgtgacgtgtagcttctatctcgtggaacaccagcgga 788

|||||

Db 66 TCACGGTTTATCAAGTTATGTTCTGTGACGCTGAGTACGTTTATCTGGACACACAGCGGA 125

|||||

QY 789 tgcatacaaaagcaagaataatgaatcaacacagagataaacaaggaaggagactttgtga 848

|||||

Db 126 TGTGATCAAAAGCAGATATGATCAACCAAGGATTAACAGAGAGGCGACTTTTGTGA 185

QY 849 taatcatcagctgactgctgttgcattcagctgttcaagtggaagatcatgagtcata 908

|||||

Db 186 TAAATCATGACAGTACGCTGATGATTCAGGCTTTCAGAGTCAAGTATCATGACTATATA 245

|||||

QY 909 taaagcttttaccatcttgctgctgagaatg 939

|||||

Db 246 TAAAGGCTTTTACCACTTGGCTGAGATG 276

|||||

RESULT 15

LOCUS BE061032 607 bp mRNA EST 09-JUN-2000

DEFINITION QV0-BT0041-061099-033-c02 BT0041 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE061032

VERSION BE061032.1 GI:8405682

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 607)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.U.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV0-BT0041-061099-033-c02&ts=1999-10-06&td=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 596.
Location/Qualifiers
1. 607
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0041"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site: 1; SmaI; Site: 2; SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 187 a 119 c 143 g 158 t

ORIGIN

Query Match 23.4%; Score 243.6; DB 162; Length 607;
Best local Similarity 75.9%; Pred. No. 4.8e-62;
Matches 369; Conservative 0; Mismatches 4; Indels 113; Gaps 1;

QY 332 acgtagtgattcttcggaggtgcgaatggttcacatcatgaaacttcgcgaggtgtgttg 391

|||||

Db 219 ACCCAGGTGATTTCTGGAGGCTGAATGCTCACATATGAAATCTCCGAGAGTGTGTTTG 278

|||||

QY 392 gcaaaagtgaagatgagcattatcccttggaaatcagtcattggagggatgatgctgc 451

```

Db 279 GCAAAAGTGAAGATGAGCATTTATCCCTTTGGAAATCAGTCATGAGGATGATGGCTG 338
Oy 452 gtgtatttggccaagttttagccaaaccactgactagtgaaagttcagatgcaaatgg 511
Db 339 GTGTATTGGCCAGTTTATAGCCATCCAACTGACCTAGTGAAGTTTCGATGCAAAATGG 398
Oy 512 aaggaanaaggaacttgaaaggaanaacatlgcgalttcgtgtgtacatgcatltg 571
Db 399 AAGGAAAAAGAACTGGAAGCAAAACCATTGC----- 431
Oy 572 caaaaatcttagctgaaggaagaatacgaaggcttggcgagctggtglaaccaatatac 631
Db 432 ----- 431
Oy 632 aaagagcagcacttggtaatatggaagalttaaccacttatgatacagtgaaacactact 691
Db 432 -----GATTTAACCACTTATGATACAGTGAACACACTACT 465
Oy 692 tggtaattgaatacaccacttgaggaacaataatcatgactcaeggttlatcaagttlatgt 751
Db 466 TGGTATTGAAATACACCACCTTGAGGACAAATATCATGACTCACGCGTTTATCAAGTTATGTT 525
Oy 752 ctggaacttgtaacttctatctctgggaacaacagccgagtgtcatcaaaagcagaataatga 811
Db 526 CTGACCTGGTACCTTCTATCTTGGAACACCAAGCGATGTCAATCCAAAGCAGAAATTAATGA 585
Oy 812 atcaac 817
Db 586 ATCCAC 591

```

Search completed: April 28, 2001, 17:51:52
 Job time: 6091 sec


```

|||||
101 ACCTGACCACTTCCCTCCGATCTCACAAAACCTGCATCCAAATGCAA 150
51 GYGLUALAALALEUALARGLEUGLYASPGLYALARGGLUSERALAPR 67
151 GGAGAGACAGCTCTGCTCGTGGAGACGGTGCAGAGAAATCTGCCCC 200
67 OTYARGGLMETVALARGTHALALEUGLYLEILEGLUGLUGLYP 84
201 CATATAGGGAATGTCGCACAGCCCTAGGAGATCATTTGAAGAGAAAGCT 250
84 HELLEULYSLEUTRPGINGLYVALTHPRCOALILETYRARGHISVALVAL 100
251 TTCATAAGCTTTGGCAGAGAGTACACCCGCCATTATACAGACCTAGAG 300
101 TYRSERGGLYARGMETVALTHRTYRGLUHSLEUARGGLVALVALPH 117
301 TATTTGGAGGTCGAATGTCACATATGACATCTCCAGAGGTTGTGTT 350
117 EGLYLYSSERGLUASPGLUHSITYRPROLEUTRPLYSSEVALILEGLYG 134
351 TGGCAAAAGTGAAAGATGAGCATTTATCCCTTGGAAATCAGTCATTGGAG 400
134 LYMETVALAGLYVALILEGLYGLINPHELEUALASNPROTTHRASPLEU 150
401 GATATATGGCTGGTGTATATGGCCAGTTTAAGCCAATCCAACTGACCTA 450
151 VALLYSVALGLINMETGLINMETGLUGLYLSARGLYLSLEUGLYLYSPR 167
451 GTGAAGGTTCAAGATGCAATGCAAGGAAAGAAAGCAAGTGAAGGAAAGC 500
167 OLEUARGPHEARGGLYVALHSHISALAPHEALALYSILEUALAGLUG 184
501 ATTTGGATTTGCTGGTGTACATCATGATTCGCAAAAATCTTACCTGAAG 550
184 LYGLYILEARGGLYLEUTRPRALAGLYTRPVALPROASNILEGLINRGLA 200
551 GAGGAAATGAGAGGCTTTGGGAGGCTGGGTACCCCAATATACAAAGACCA 600
201 ALALEUVALASNMETGLYASPLEUTHRTYRASPTRHVALYSHISTY 217
601 GCACGTGGATATGAGAGATTTAACCACTATGATACAGTCAAAACACTA 650
217 TLEUVALLEUASNTHRPROLEUGLYASPNLLEWETHRHISGLYLEUS 234
651 CTGTGATATGAATACACCACTTGAGGACAATATCATGACTCAGGTTTAT 700
234 ERSERLEUCYSSERGLYLEUVALALASERILEUGLYTHPRALASP 250
701 CAAGTTATGTTCTGACTGTGACTTCTATTTCTGGAAACACACCCGAT 750
251 VALILEYSSERARGILEMETASNGINPROARGASPLYSGLINGLYARGI 267
751 GTCATCAAAACAGAAATATGATCAACACGAGATAACAGAGAGAGGG 800
267 YLEULEUTRILYSSERSETRHASPYSLEUILLEGLNALVALINGLYG 284
801 ACTTTGTATATAATCATGACTGACTGCTGATTCAGGCTGTCAAGGTG 850
284 LUGLYPHEMETSERLEUTRILYSGLYPHELEUPROSETRIPLEUARGMET 300
851 AAGGATTCATAGTCTATATAAGGCTTTTACCACTTGGCTGGAATG 900
301 THPROTRPSERMETVALPHETRIPLEUTHRTYRGLULYSILEARGGLIME 317
901 ACCCCTTGCTCAATGCTGTCTGCTTACTATGAAAAAATCAGAGAGAT 950
317 TSEGLYVALSERPROPE 323
951 GAGTGGAGTCAATCATTT 969
seq_name: gb_ro2:RNO300162

```

```

seq_documentation_block:
LOCUS      RNO300162      2678 bp      mRNA      06-JAN-2001
DEFINITION Rattus norvegicus mRNA for uncoupling protein UCP-4 (ucp-4 gene),
isoform a.
ACCESSION AJ300162.1 GI:12055539
VERSION    AJ300162.1
KEYWORDS   Isoform a; ucp-4 gene; uncoupling protein UCP-4.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 2678)
AUTHORS   Alberati-Giani,D., Gatti,S., Rial,E., Danel,F., Hauser,N.,
            Bubendorff,C. and Barfai,T.
TITLE      Three different isoforms of UCP-4 are expressed in rat preoptic
            anterior hypothalamus
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 2678)
AUTHORS   Alberati-Giani,D.
TITLE      Direct Submission
JOURNAL    Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division
            preclinical Research, CNS, F. Hoffmann-La Roche Ltd.,
            Grenzachstrasse 124, CH 4070 Basel, SWITZERLAND
FEATURES   Location/Qualifiers
            source
            1..2678
            /organism="Rattus norvegicus"
            /strain="Sprague Dawley"
            /db_xref="taxon:10116"
            /sex="male"
            /russue_type="brain, preoptic anterior hypothalamus"
            /dev_stage="6 weeks"
            /russue_lib="UNI-ZAP custom cDNA library"
            /country="Switzerland"
            52..1020
            /gene="ucp-4"
            52..1020
            /gene="ucp-4"
            /function="uncouples respiration in eukaryotic cells"
            /codon_start=1
            /product="uncoupling protein UCP-4, isoform a"
            /protein_id="CAC20898.1"
            /db_xref="GI:12055540"
            /translation="MPPEEESLQPLTQWRPRTSKFLSGCAATVAELATFPDLTKT
            RLOMGEAALAKLGDGAMESAPYRGMMRTALGILOBEGLKLMQGVPAIVRYVVG
            GRMVTYEHLEEVYFGKSEDEHYPLKMSVYCGMAGVYIGCFALNPPLVYVQMGESKR
            RLEKPLREGVNHAHAKITLDESGTGLMAGWIPNORALVNMGLDITVDYKHLV
            LNTALDENIATHQLSSLCGLVASILGTADVILKSRIMQPRDKRGRLYKSSDTCV
            IDAVDGEGLSLYKGLFPLSWLRMTFPMVSMVFWLTYEKIRQLSGVSP"
BASE COUNT      683 a      674 c      676 g      645 t
ORIGIN
alignment_scores:
Quality: 1603.50      Length: 323
Ratio: 5.074      Gaps: 1
Percent Similarity: 97.833      Percent Identity: 94.118
alignment_block:
US-09-397-342-1 x RNO300162
Align seg 1/1 to: RNO300162 from: 1 to: 2678
1 MetSerValProGluGluGluGluArgLeuLeuProLeuThrGlnArgTr 17
|||||
52 ATGCCTTCCCTCGAGAGAGAGTCG...CTACAGCCACTCAGCAGAGATG 98
17 ProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
|||||
99 GCCCGGACAGCAAGTCTACTGTCTCGGCGCCGCCCAACCGTGGCGG 148
17 LeuLeuAlaThrPheProLeuLeuAspLeuThrLysThrArgLeuGlnMetGln 50
|||||

```


257 AGCTAGCAACCTTTCCCTGATCTCACAACCAACCGCATCCGATGCAA 306
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAla 67
|||||
307 GGAGAAAGCTCCCTTCTAAGTTGGAGATGTGTGCATGTGAATCCGCC 356
67 oTyraArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGly 84
|||||
357 TTACAGGGCATATGTCGACGGCCCTGGGATGTTCACAGAGGAAGGCT 406
84 heLeuLysLeuThrPrgInclGlyValThrProAlaIleTyraArgHisVal 100
|||||
407 TCCTAAAGTTGGGCAAGAGTGCACGCCCATTTACAGACACGTAAGT 456
101 TyrSerGlyGlyArgMetValThrTyrgLuhHisLeuArgGluValAla 117
|||||
457 TACTGTGAGAGTCCGATGTCTACCTACGAACATCTCCGGGAAGTGTGT 506
117 eGlyLysSerGluAspGluHisTyrgProLeuThrPlySerValIleGly 134
|||||
507 TGGCAAAAGTGAAGATGAACATTATCCCTGTGGAAGTCCGTCATTGG 556
134 lYmetMetAlaGlyValIleGlyLuhPheLeuAlaAsnProThrAspLeu 150
|||||
557 GGATGATGGCTGGCGCATCGACACACTTTTACCCACACCCACTGACCT 606
151 ValLysValGluMetGluMetGluGlyLysArgLysLeuGluGlySer 167
|||||
607 GTGAAGGTCCAGATGCCAATGGAAGGAAACGTAGACTGGAAGGAAC 656
167 oLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
|||||
657 GTTGAGATTCCGGGAGTACATCATCGTTTGCAAAATCTTACGTGAAG 706
184 lYgLyIleArgGlyLeuThrPalaglyTyrgValProAsnIleGluAla 200
|||||
707 GAGGATCGGTGGGCTTTGGCAGCGTGTATGCCAATATTCAGAGACA 756
201 AlAlaLeuAlaMetGlyAspLeuThrTyrgAspThrValLysHisTy 217
|||||
757 GCACCTGTGACATGGAGATTAAACCACTTACGACACAGTGAAGACAT 806
217 rLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisGlyLeu 234
|||||
807 CCTGGATTGAACACACGACCTTGAAGACAATATCGCCACTCAGCGCT 856
234 eTserLeuLysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
|||||
857 CCAGTTGTGTTCTGACGTGCGCTTCTATTTCTGGGAACACACGCGCAT 906
251 ValIleLysSerArgIleMetAsnGluProArgAspLysGluGlyArg 267
|||||
907 GTTCATCAAAAGCCGAATATGAACCACTCGACACAAACAGGAAGAG 956
267 vLeuLeuTyrgLysSerSerThrAspCysLeuIleGluAlaValGluGly 284
|||||
957 GCTTTTGTATTAAGTCACTCACTGACGCGTGTATCCAGCTGTCAAGAG 1006
284 lUglYpMetSerLeuTyrgLysGlyPheLeuProSerThrPleuArgMet 300
|||||
1007 AAGGCTTCTCGACCGCTATTAAGCTTCTACCCCTCCGCTGAGCAATG 1056
seq_name: gb_r02:RNO300163
seq_documentation_block:
LOCUS RNO300163 2620 bp mRNA ROD 06-JAN-2001
DEFINITION Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)
isoform b.
ACCESSION AJ300163 GI:12055541
VERSION AJ300163.1
KEYWORDS isoform b; Ucp-4 gene; uncoupling protein UCP-4.
SOURCE Norway rat.
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 2620)
Alberati-Giani,D., Gatti,S., Rial,E., Danel,F., Hauser,N.,
Bubendorf,C. and Barfai,T.
TITLE
Three different isoforms of UCP-4 are expressed in rat preoptic
anterior hypothalamus
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2620)
Alberati-Giani,D.
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (02-NOV-2000) Alberati-Giani D., Pharma Division
Preclinical Research, CNS, F. Hoffmann-La Roche Ltd.,
Grenzacherstrasse 124, CH 4070 Basel, SWITZERLAND
FEATURES
source
1..2620
/organism="Rattus norvegicus"
/strain="Sprague Dawley"
/db_xref="taxon:10116"
/sex="male"
/issue_type="brain, preoptic anterior hypothalamus"
/dev_stage="6 weeks"
/tissue="lb="UNT-ZAP custom cDNA library"
/country="Switzerland"
/gene="ucp-4"
139..1173
/gene="ucp-4"
139..1173
/function="uncouples respiration in eukaryotic cells"
/codon_start=1
/product="uncoupling protein UCP-4, isoform b"
/protein_id="CAC20899.1"
/db_xref="GI:12055542"
/translation="MPPEEESLQPLNRPRTSKFLISGCAATVAELATPELDLTK
RLMOGREALAKISDGSAPRGCMRTALGYOEGFLKMGVATPAIRHYVYG
GRVATYHLREYVYVGSDEHYPLMKSVIGMAAGVIGOTLANTDLVKOMQEGH
RLEKGPLRFKRVHFAKILAEGLIRGLMAGWIPNIGRALVNNGLDITVDYKHYL
LNLLEDNIAITHGLSLCSGIVASILCTPADVIRSRIMNPPRODGGRLYKSSDGV
IOAVOGGFLSLYGFPLSWLRLMEDTLHRCDLQGATLYPLETGVLSDSFYLRAL
CLVOSHETNNH"
BASE COUNT 657 a 661 c 679 g 623 t
ORIGIN
alignment_scores:
Quality: 1480.50 Length: 300
Ratio: 5.053 Gaps: 1
Percent Similarity: 97.667 Percent Identity: 94.333
alignment_block:
US-09-397-342-1 x RNO300163 ..
Align seg 1/1 to: RNO300163 from: 1 to: 2620
1 MetSerValProGluGluGluGluArgLeuLeuProLeuThrGlnArgTr 17
|||||
139 ATGCGCTTCCCTGAGAGAGATCG...CTACAGCCCACTCACAGAGATG 185
17 pProArgAlaSerLysPheLeuLeuSerGlyLysAlaAlaThrValAlaG 34
|||||
186 GCCCGGACCGACGACAGTCTCTACTGTCCGCTGCGCGCAACCGTGGCG 235
34 lUeUAlaIaThrPheProLeuAspLeuThrLysThrArgGluMetGln 50
|||||
236 AGCTAGCAACCTTTCCCTGCATCTCACAAAACCGCATCCAGATGCAA 285
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAla 67
|||||
286 GGAGAAAGTCCCTTCTAAGTTGGAGATGTGTGCATGTGAATCCGCC 335
67 oTyraArgGlyMetValArgThrAlaLeuGlyIleIleGluGluGly 84
|||||


```

336 TTACAGGGGCGATGATCCGACGGCCCTGGGATGTCCAGAGGAGGCT 385
84 heLeuLysLeuThrPglnglyValThrProAlaIleTyrArgHisValVal 100
386 TCCAAAGTTGTGGCAGAGGAGGAGCGCCGCTTATACAGACACGCTAGTG 435
101 TTTSerGlyGlyArgMetValThrTyrGluHisLeuArgGluValValPh 117
436 TACTCTGGAGGTCGGATGTCACCTACCAACATCTCCGGGAAGTTGTGT 485
117 eGlyLysSerGluAspGluHisTyrProLeuThrPlysSerValIleGly 134
486 TGGCAAAAGTGAAGATGAACATTAATCCCTCGAAGTCCGTCATTTGAG 535
134 lymeMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
536 GGATGATGCGTGGCGTCATCGACAGTTTAAAGCCACCCACCTGACCTC 585
151 ValLysValGlnMetGlnMetGluGlyLysArgLysLeuGluGlyLysPr 167
586 GTGAAGGTCGAGATGCAAAATGGAAGGGAACCTGACCTGGAAGGGAACC 635
167 oLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
636 CTTGAGATTCGGGGAGTACATCATGCTTGGCAAAATCTTACCTGAG 685
184 lYgylIleArgGlyLeuThrPAlaGlyTTPValProAsnIleGlnArgAla 200
686 GAGGATATCGTGCGCTTGGCGAGGCTGTGATACCAATATTCAGAGACA 735
201 AlaLeuValAsnMetGlyAspLeuThrThrTyrAspThrValLysHisTyr 217
736 GCACCTTGAGACATGAGAGATTTAAACCATACGACACAGTGAACACTA 785
217 lLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisGlyLeu 234
786 CCGTGTATTGAACACAGCACTTGAGACAAATATCCCACTCACGCTTAT 835
234 erSerLeuGlySerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
836 CCAGTTGTGTGTGACTGTGGCTTCTATTCGTGGAAACACACCGCAT 885
251 ValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArgG 267
886 GTCATCAAAAGCCGAAATATGACCAACCTCGAGCAAAACAGAGAGAG 935
267 yLeuLeuThrLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
936 GCTTTTGTATTAAGTCATCACTGACTGCGTATCCAGGCTGTTCAAGAG 985
284 lGlyPheMetSerLeuThrLysGlyPheLeuProSerThrLeuArgMet 300
986 AAGGCTTCTGAGCGCTGTATAAAGGCTTTCAACCTCGCTGAGGAATG 1035
seq_name: gb_pat1:AX061217
seq_documentation_block:
LOCUS AX061217 1726 bp DNA PAT 22-JAN-2001
DEFINITION Sequence 64 from Patent WO0078953.
ACCESSION AX061217
VERSION AX061217.1 GI:12406353
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1726)
REFERENCE
AUTHORS Lal, P., Yang, J., Yue, H., Hillman, J. L., Tang, Y. T., Bandman, O.,
Burford, N., Baughn, M. R., Azimzal, Y., Lu, D. A., Au-Young, J., and
Patterson, C.
Human transport proteins
Patent: WO 0078953-A 64 28-DEC-2000;
Incyte Genomics, Inc. (US)

```

```

FEATURES
    source
        location/Qualifiers
            1..1726
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 552 a 302 c 399 g 473 t
ORIGIN
alignment_scores:
    Quality: 1186.50 Length: 244
    Ratio: 5.181 Gaps: 1
    Percent Similarity: 93.852 Percent Identity: 92.623
alignment_block:
US-09-397-342-1 x AX061217
Align seg 1/1 to: AX061217 from: 1 to: 1726
95 lLeTyrArgHisVal..... 99
397 ATACATTAAGCATATATGATATATAGCCAAAGTTACCTTTTAAATGATCT 446
100 .....ValTyrSerGlyGlyArgMetValThrTyrGluHisLeu 113
447 TTTTACCACGATGTATTCGAGAGTCGAATGCTACATATGAACATCTCC 496
113 rGlyValValPheGlyLysSerGluAspGluHisTyrProLeuThrPlys 129
497 GAGAGGTTGTGTTTGCAAAAGTGAAGATGACATTAATCCCTTGGAAA 546
130 SerValIleGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAs 146
547 TCAGTCATTGGAGGAGATGCTGCTGTATTGCGCAGTTTATGCCAA 596
146 nProThrAspLeuValLysValGlnMetGlnMetGluGlyLysArgLysL 163
597 TCCAACTGACCTAGTGAAGGTTCAAGTGCATGCAATGGAAGGAAAAAGAAAC 646
163 euGlnGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 179
647 TGGAGGAAACCAATTCGATTCGTGTACATCATGCAATTCACAAA 696
180 lLeuAlaGluGlyGlyLeuArgGlyLeuThrPAlaGlyTTPValProAs 196
697 ATCTTAAGCTGAAGAGAGATATGAGAGGCTTGGGAGGCTGGTACCCAA 746
747 TATACAAAGAGCAGACACTGCTGAATATGAGAGATTTAACCACTTATGATA 796
213 hValLysHisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMet 229
797 CAGTAAACACTACTTGTGATTAATACACCACTTGAGGACAAATATCATG 846
230 ThrHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuG 246
847 ACTCAAGGTTTATCAAGTTATGTCTGACTGCTGAGCTTATTCCTGGG 896
246 yThrProAlaAspValIleLysSerArgIleMetAsnGlnProArgAspL 263
897 AACACAGCGCGATGTCATCAAAAGAGATATATGATCAACACAGATA 946
947 AACACAGAGAGGAGCTTTTGTATTAATCATCAGCTGCTTGTGATGAG 996
263 yGlnGlyArgGlyLeuLeuThrLysSerSerThrAspCysLeuIleGln 279
280 AlaValGlnGlyGlyPheMetSerLeuThrLysGlyPheLeuProse 296
997 GCTGTTCAAGGAGGAGATTCATGAGCTATATTAAGGCTTTTATCAATC 1046
296 rThrPLeuArgMetThrProTyrSerMetValPheThrPLeuThrTyrGlu 313
1047 TTGGCTGAAGATGACCCCTTGCTCAATGATGTCTGCTTACTTATGAAA 1096

```

313 ys1learglunetSerglyValSerProPhe 323
 1097 AAATCAGACAGATGATGAGTGCATGTCATTT 1128
 seq_name: gb_htg2:AC009216

seq_documentation_block:
 LOCUS AC009216 169457 bp DNA HTG 03-DEC-1999
 DEFINITION Drosophila melanogaster chromosome X clone BACR38423 (D900) RPCI-98
 38 H.23 map 16D-16f strain Y; cn bw sp, *** SEQUENCING IN PROGRESS
 *** 118 unordered pieces.

ACCESSION AC009216 GI:6513905
 VERSION AC009216.8
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
 1 (bases 1 to 169457)
 Burenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
 Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
 Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
 Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
 Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
 Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,
 Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.

Sequencing of Drosophila melanogaster
 Unpublished
 2 (bases 1 to 169457)
 Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
 Burenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
 Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
 Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
 Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Macda,P.,
 Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
 Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
 Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
 Rubin,G.M.

TITLE Direct Submission
 JOURNAL Submitted (06-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 COMMENT On Dec 3, 1999 this sequence version replaced gi:6478893.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our sequence
 archive web site (<http://www.fruitfly.org/sequence/>) or send email
 to hdg@fruitfly.berkeley.edu. All contigs in this submission meet
 the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 118 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 489: contig of 489 bp in length
 * 490 569: gap of unknown length
 * 570 1126: contig of 557 bp in length
 * 1127 1206: gap of unknown length
 * 1207 1561: contig of 355 bp in length
 * 1562 1641: gap of unknown length
 * 1642 2169: contig of 528 bp in length
 * 2170 2249: gap of unknown length
 * 2250 2833: contig of 584 bp in length
 * 2834 2913: gap of unknown length
 * 2914 3802: contig of 889 bp in length
 * 3803 3882: gap of unknown length
 * 3883 4828: contig of 946 bp in length
 * 4829 4908: gap of unknown length
 * 4909 5573: contig of 665 bp in length

* 5574 5653: gap of unknown length
 * 5654 6272: contig of 619 bp in length
 * 6273 6352: gap of unknown length
 * 6353 7267: contig of 915 bp in length
 * 7268 7347: gap of unknown length
 * 7348 7900: contig of 553 bp in length
 * 7901 7980: gap of unknown length
 * 7981 9022: contig of 1042 bp in length
 * 9023 9102: gap of unknown length
 * 9103 9692: contig of 590 bp in length
 * 9693 9772: gap of unknown length
 * 9773 10394: contig of 622 bp in length
 * 10395 10474: gap of unknown length
 * 10475 10877: contig of 403 bp in length
 * 10878 10957: gap of unknown length
 * 10958 11712: contig of 755 bp in length
 * 11713 11792: gap of unknown length
 * 11793 12174: contig of 382 bp in length
 * 12175 12254: gap of unknown length
 * 12255 12898: contig of 644 bp in length
 * 12899 14013: contig of 1035 bp in length
 * 14014 14094: gap of unknown length
 * 14094 15066: contig of 972 bp in length
 * 15066 15146: gap of unknown length
 * 15146 15914: contig of 769 bp in length
 * 15915 16419: gap of unknown length
 * 15995 16419: contig of 425 bp in length
 * 16420 16499: gap of unknown length
 * 16500 17461: contig of 962 bp in length
 * 17462 17541: gap of unknown length
 * 17542 18098: contig of 557 bp in length
 * 18099 18178: gap of unknown length
 * 18179 18988: contig of 810 bp in length
 * 18989 19069: gap of unknown length
 * 19069 20212: contig of 1144 bp in length
 * 20213 20292: gap of unknown length
 * 20293 20817: contig of 525 bp in length
 * 20818 22197: gap of unknown length
 * 20898 22277: gap of unknown length
 * 22198 22278: gap of unknown length
 * 22278 22438: contig of 1161 bp in length
 * 22439 23518: gap of unknown length
 * 23519 24889: contig of 1371 bp in length
 * 24890 24969: gap of unknown length
 * 24970 25810: contig of 841 bp in length
 * 25811 25890: gap of unknown length
 * 25891 27149: contig of 1259 bp in length
 * 27150 27229: gap of unknown length
 * 27230 28261: contig of 1032 bp in length
 * 28262 28341: gap of unknown length
 * 28342 29712: contig of 1371 bp in length
 * 29713 29792: gap of unknown length
 * 29793 31385: contig of 1593 bp in length
 * 31386 31466: gap of unknown length
 * 31466 32803: contig of 1338 bp in length
 * 32804 32883: gap of unknown length
 * 32884 34706: contig of 1823 bp in length
 * 34707 34786: gap of unknown length
 * 34787 36436: contig of 1650 bp in length
 * 36437 36516: gap of unknown length
 * 36517 38395: contig of 1879 bp in length
 * 38396 38475: gap of unknown length
 * 38476 39602: contig of 1127 bp in length
 * 39603 39682: gap of unknown length
 * 39683 41589: contig of 1907 bp in length
 * 41590 41669: gap of unknown length
 * 41670 43475: contig of 1806 bp in length
 * 43476 43555: gap of unknown length
 * 43556 45795: contig of 2240 bp in length
 * 45796 45875: gap of unknown length
 * 45876 48574: gap of unknown length
 * 48575 48654: gap of unknown length

```

* 48655 50881: contig of 2227 bp in length
* 50882 50961: gap of unknown length
* 50962 52818: contig of 1857 bp in length
* 52819 52898: gap of unknown length
* 52899 55256: contig of 2358 bp in length
* 55257 55336: gap of unknown length
* 55337 59040: contig of 3704 bp in length
* 59041 59120: gap of unknown length
* 59121 62581: contig of 3461 bp in length
* 62582 62661: gap of unknown length
* 62662 62665: contig of 3604 bp in length
* 62666 66346: gap of unknown length
* 66347 71526: contig of 5181 bp in length
* 71527 71606: gap of unknown length
* 71607 76897: contig of 5291 bp in length
* 76898 76977: gap of unknown length
* 76978 83732: contig of 6755 bp in length
* 83733 83812: gap of unknown length
* 83813 89149: contig of 5337 bp in length
* 89150 89229: gap of unknown length
* 89230 97652: contig of 8423 bp in length
* 97653 97732: gap of unknown length
* 97733 106245: contig of 8512 bp in length
* 106246 106324: gap of unknown length
* 106325 116692: contig of 10368 bp in length
* 116693 116772: gap of unknown length
* 116773 129843: contig of 13071 bp in length
* 129844 129923: gap of unknown length
* 129924 130419: contig of 496 bp in length
* 130420 130499: gap of unknown length
* 130500 131145: contig of 646 bp in length
* 131146 131225: gap of unknown length
* 131226 131840: contig of 615 bp in length
* 131841 131920: gap of unknown length
* 131921 132349: contig of 629 bp in length
* 132350 132629: gap of unknown length
* 132630 133129: contig of 500 bp in length
* 133130 133209: gap of unknown length
* 133210 133458: contig of 249 bp in length
* 133459 133538: gap of unknown length
* 133539 134232: contig of 714 bp in length
* 134233 134352: gap of unknown length
* 134353 134855: contig of 523 bp in length
* 134856 134935: gap of unknown length
* 134936 135396: contig of 461 bp in length
* 135397 135476: gap of unknown length
* 135477 135988: contig of 512 bp in length
* 135989 136068: gap of unknown length
* 136069 136855: contig of 787 bp in length
* 136856 136935: gap of unknown length
* 136936 137763: contig of 828 bp in length
* 137764 137843: gap of unknown length
* 137844 138304: contig of 461 bp in length
* 138305 138384: gap of unknown length
* 138385 138905: contig of 521 bp in length
* 138906 138985: gap of unknown length
* 138986 139626: contig of 641 bp in length
* 139627 139706: gap of unknown length
* 139707 140425: contig of 719 bp in length
* 140426 141245: gap of unknown length
* 141246 141325: contig of 740 bp in length

```

alignment_scores:

Quality: 753.50 Length: 323
 Ratio: 3.456 Gaps: 4
 Percent Similarity: 67.492 Percent Identity: 46.440

alignment_block:

US-09-397-342-1 x AC009216

Align seg 1/1 to: AC009216 from: 1 to: 169457

```

68 TyrArgGlyMetValArgThrAlaLeuGlyIleIleGlyGlyGlyPhe 84
|||||.....|
122104 TACCCGGGCAATGCTGGCCACCGCCTTGGGATGCGGCTGAGAGAGGCG 122153
|||||.....|
84 LeuLeuLeuTrpGlnGlyValThrProAlaIleTyrArgHisValValT 101
|||||.....|
122154 CCTGAAGCTGTGGCAGGGCGTAACGCCGGCGCTTACCGACACGCTGCT 122203
|||||.....|
101 Trp..... 101
122204 ATAGGTGGGCATTCCTCAAACTCTCATTTAGCATATGATCACCGCCA 122253
|||||.....|
101 ..... 101
122254 CAGCCAACTGCGCCGACGACGCTGTAGCCATCAGCATTAAGTACGACAG 122303
|||||.....|
102 ..... 102
122304 TTCTAACCCAACTTCTCCCTGCTCTCCATGCTCCATCCGACAGGGTGTGA 122353
|||||.....|
105 rgMetValThrTyrGlnHisLeuArgGluValAlaPheGlyLysSerGlu 121
|||||.....|
122354 GGAATCGACGCTACGACCTGATGCGCAAG...GAGTTCAACAGAAACGCG 122400
|||||.....|
122401 ACCGAGGCTGCGCGGTTGGAAAGTCGGCGCTGCGGCTGACGACGCGG 122450
|||||.....|
138 YValIleGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGln 155
|||||.....|
122451 AGCCCTTGCACAGTGGCTTGGCTGCGCGGCTGACGCTGCAAGGTGCAAA 122500
|||||.....|
155 etGlnMetGlnGlyLysArgLysLeuGlyLysProLeuArgPheArg 171
|||||.....|
122501 TCCAATGAGGCGCGACGATGCTGATGGGCGCCACGAGGGGTGAC 122550
|||||.....|
172 GlyValHisHisAlaPheAlaLysIleLeuAlaGlnGlyGlyIleArgG 188
|||||.....|
122551 TCCGCTGGCCATGCTTCCGCCAGATGCTGCAGCTGGCGGAATTAAGGG 122600
|||||.....|
188 YLeuTrpAlaGlyTrpValProAsnIleGlnArgAlaAlaLeuValAsn 205
|||||.....|
122601 CCTGTGGAAGGACAGCATCCGGAATGTCACGAGCGCGTGTGTCATC 122650
|||||.....|
205 etGlyAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsn 221
|||||.....|
122651 TGGGGACCTTACACCATATACGACATCAAGCATCATGATCATGATCGC 122700
|||||.....|
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerLeuGlySer 238
|||||.....|
122701 CTGCAGATGCCCGACTGTACACACATGACATGCTGGCTCCGTTGTGTC 122750
|||||.....|
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerA 255
|||||.....|
122751 CGGATTCGTGGCAGGATCATGAGGACCGCAGCTGATGTGTAAGACGC 122800
|||||.....|
255 rGlyLeuMetSerGlnProArgAspLysGlnGlyArg..... 266
|||||.....|
122801 GCATCATGAAACAGCCACGACGAGAAATGCGCGTAGTAGCAACAATGCG 122850
|||||.....|
267 ..... 267
122851 CGCATTTACTATTGGAACGCTTAATCACTATGTTATTCGCGAGGGGC 122900
|||||.....|
268 LeuLeuTyrLysSerSerThrAspLysLeuIleGlnAlaValGlnGlyG 284
|||||.....|
122901 CTGCTCATTCGCGGATCGGTGACGCTCGGACAGACAGAGCGTTTGAAGA 122950
|||||.....|
284 uGlyPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMetT 301
|||||.....|
122951 GGGCTTGTGGCGCTGTACAAAGGCTTTCCTCGCTGCTGATCAAGTGG 123000
|||||.....|
301 hrProTrpSerMetValPheTrpLeuThrTyrGlnLysIleArgGlnMet 317
|||||.....|

```

123001 CGCCGTGCTGCTACCTTCTGCTGCTTCAACAGATCCGACAGATG 123050
318 SerGIYalSerProPhe 323
123051 ATCGAGCCTCCGCGCTAC 123068

seq_name: gb_hyt4:AC012162

seq_documentation_block:
LOCUS AC012162 191504 bp DNA HTG 09-FEB-2000
DEFINITION Drosophila melanogaster chromosome X locus BACR01N10 (D1115)
PPT-98 01 N.10 map 16D-16F strain Y; cn bw sp, *** SEQUENCING IN
PROGRES *** 142 unordered pieces.

ACCESSION AC012162
VERSION AC012162.9 GI:6957580
KEYWORDS HTG; HTGS_PHASE1.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 191504)

REFERENCE
AUTHORS Celniker,S.E., Aabayani,A., Arcalata,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hindle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirskas,R.R., Man,K.H., Webster,D.,
Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 191504)

TITLE
JOURNAL
REFERENCE
AUTHORS Celniker,S.E., Aabayani,A., Arcalata,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequelira,A., Sethi,H., Shlr,E.,
Svirskas,R.R., Man,K.H., Weinburg,T., Zhang,R., Zieren,L.L. and
Rubin,G.M.

TITLE
JOURNAL
COMMENT Direct Submission
Submitted (21-Oct-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 11, 2000 this sequence version replaced gi:6838825.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 142 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 456: contig of 456 bp in length
* 457 536: gap of unknown length
* 537 1004: contig of 468 bp in length
* 1005 1084: gap of unknown length
* 1085 1958: contig of 874 bp in length
* 1959 2038: gap of unknown length
* 2039 2248: contig of 210 bp in length
* 2249 2328: gap of unknown length
* 2329 2858: contig of 530 bp in length
* 2859 2938: gap of unknown length
* 2939 3325: contig of 387 bp in length
* 3326 3405: gap of unknown length
* 3406 4165: contig of 760 bp in length

* 4166 4245: gap of unknown length
* 4246 4716: contig of 471 bp in length
* 4717 4796: gap of unknown length
* 4797 5373: contig of 577 bp in length
* 5374 5453: gap of unknown length
* 5454 6050: contig of 597 bp in length
* 6051 6130: gap of unknown length
* 6131 6720: contig of 590 bp in length
* 6721 6800: gap of unknown length
* 6801 7461: contig of 661 bp in length
* 7462 7542: gap of unknown length
* 7543 8104: contig of 563 bp in length
* 8105 8184: gap of unknown length
* 8185 8505: contig of 321 bp in length
* 8506 8585: gap of unknown length
* 8586 9227: contig of 642 bp in length
* 9228 9307: gap of unknown length
* 9308 10091: contig of 784 bp in length
* 10092 10171: gap of unknown length
* 10172 11099: contig of 928 bp in length
* 11099 11179: gap of unknown length
* 11180 11747: contig of 568 bp in length
* 11748 11827: gap of unknown length
* 11828 12767: contig of 940 bp in length
* 12768 12847: gap of unknown length
* 12848 13398: contig of 551 bp in length
* 13399 13478: gap of unknown length
* 13479 14007: contig of 529 bp in length
* 14008 14087: gap of unknown length
* 14088 14867: contig of 780 bp in length
* 14868 14947: gap of unknown length
* 14948 15562: contig of 615 bp in length
* 15563 15642: gap of unknown length
* 15643 16859: contig of 1217 bp in length
* 16860 16939: gap of unknown length
* 16940 17986: contig of 1047 bp in length
* 17987 18066: gap of unknown length
* 18067 19341: contig of 1275 bp in length
* 19342 20926: contig of 1505 bp in length
* 20927 21006: gap of unknown length
* 21007 22153: contig of 1147 bp in length
* 22154 22233: gap of unknown length
* 22234 23236: contig of 1003 bp in length
* 23237 23316: gap of unknown length
* 23317 24607: contig of 1291 bp in length
* 24608 24687: gap of unknown length
* 24688 25620: contig of 933 bp in length
* 25621 25700: gap of unknown length
* 25701 26300: contig of 600 bp in length
* 26301 26380: gap of unknown length
* 26381 27488: contig of 1108 bp in length
* 27489 27568: gap of unknown length
* 27569 28215: contig of 647 bp in length
* 28216 28295: gap of unknown length
* 28296 29440: contig of 1145 bp in length
* 29441 29520: gap of unknown length
* 29521 30587: contig of 1067 bp in length
* 30588 30667: gap of unknown length
* 30668 32042: contig of 1375 bp in length
* 32043 32122: gap of unknown length
* 32123 33740: contig of 1618 bp in length
* 33741 33820: gap of unknown length
* 33821 35375: contig of 1555 bp in length
* 35376 35455: gap of unknown length
* 35456 36258: contig of 803 bp in length
* 36259 36338: gap of unknown length
* 36339 38003: contig of 1665 bp in length
* 38004 38084: gap of unknown length
* 38085 39509: contig of 1426 bp in length
* 39510 39589: gap of unknown length
* 39590 40719: contig of 1130 bp in length
* 40719 40799: gap of unknown length

RLREFPNLEOASVVOFGFESIEALKRFCKAKPEKSKPIIISGSGYKSSPTSDNSCS
SSPTGNGSGFIIPLGSSSSMSNLISDSTSPSSSSSVNTIANGROOQOQOQOQOQO
PDVSGGPPERRPTKOEIYERFKOIEGHIDIKVITINENPFISSGDTPTSLKECC
RYNAMIICAQVANKARIQOILKLTISDREFTOLYVKRSGKACALNLSLDIYINMP
DKRGETPLHEFAKNGHVAWEVLVSPECKSLRNHEKEKEPIICLRNANATHYTIK
KLELIYDHPHEVLRSSQNTLPKVGOPEFKDPIIDPLNQHKADEGSLVDAIHALA
GMSREKAMENFYRRMKTTPRVSNMYPAGSPFSPKVPYKSIIFDPSASNSPVH
SGRVLFSPLAATSSPKTKVPMGTNCEENNINNVAVPYLEPAPDIRMKRDLF
MAYRNNSFDSPLADDSOILDMSLSRNLASLNSFRERHKNKDIEKGLYEVROL
AROELEHRETFIDLSFDITGTLARLEAFYEKTEQOAKSEITVWNFHNOYF
DSMAGEOQO
SINDTILCELRKLSLIVSFKDARFISVDESKVSRILAHVAASYTHSOEVSAMRL
OILOMLRSLROLADREGRHOHLGCVCASILLMLBOAPTSVAHLBDTLKTEELCAAM
ETEOCCACIMDANLSRKTSSRRKTKSLRAAAVVOGOLODTSGSSSDEHEMERT
STIGASRYVASAKDAMRROOSDDEYDSDPOVTEPCTNTLTLEYGSSSEDEHEMERT
PROSLSPGISMDLEPRYELFTGNENPTKRDIDYLAHSVNDIKETLHVIAWKAME
STSCAEMNLNVAVOKPEPWYSGTSSSHNSQPLHFKRLIATPKLAVVGSRGSPVLT
APVTPLRLATPSAASIOVASETNGSVGTAVTPASPIISFALITRAATOSFOTPLMKVR
GLFSQYRDORSYNEGDTPLGNRN"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28092))
/gene="CG8465"
/note="CG8465 gene product [alt 2]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAFA8736.1"
/db_xref="GI:7293358"
/translation="MSTEGVYIPTSKAGEGYSOCIGSIANAVINRPNPASGSAS
VASGSPGSAASVOTGNADGSAATKEDPDPPSPMLIFTEKSKSLDILRHKEA
RLREFPNLEOASVVOFGFESIEALKRFCKAKPEKSKPIIISGSGYKSSPTSDNSCS
SSPTGNGSGFIIPLGSSSSMSNLISDSTSPSSSSSVNTIANGROOQOQOQOQOQO
PDVSGGPPERRPTKOEIYERFKOIEGHIDIKVITINENPFISSGDTPTSLKECC
RYNAMIICAQVANKARIQOILKLTISDREFTOLYVKRSGKACALNLSLDIYINMP
DKRGETPLHEFAKNGHVAWEVLVSPECKSLRNHEKEKEPIICLRNANATHYTIK
KLELIYDHPHEVLRSSQNTLPKVGOPEFKDPIIDPLNQHKADEGSLVDAIHALA
GMSREKAMENFYRRMKTTPRVSNMYPAGSPFSPKVPYKSIIFDPSASNSPVH
SGRVLFSPLAATSSPKTKVPMGTNCEENNINNVAVPYLEPAPDIRMKRDLF
MAYRNNSFDSPLADDSOILDMSLSRNLASLNSFRERHKNKDIEKGLYEVROL
AROELEHRETFIDLSFDITGTLARLEAFYEKTEQOAKSEITVWNFHNOYF
DSMAGEOQO
SINDTILCELRKLSLIVSFKDARFISVDESKVSRILAHVAASYTHSOEVSAMRL
OILOMLRSLROLADREGRHOHLGCVCASILLMLBOAPTSVAHLBDTLKTEELCAAM
ETEOCCACIMDANLSRKTSSRRKTKSLRAAAVVOGOLODTSGSSSDEHEMERT
STIGASRYVASAKDAMRROOSDDEYDSDPOVTEPCTNTLTLEYGSSSEDEHEMERT
PROSLSPGISMDLEPRYELFTGNENPTKRDIDYLAHSVNDIKETLHVIAWKAME
STSCAEMNLNVAVOKPEPWYSGTSSSHNSQPLHFKRLIATPKLAVVGSRGSPVLT
APVTPLRLATPSAASIOVASETNGSVGTAVTPASPIISFALITRAATOSFOTPLMKVR
GLFSQYRDORSYNEGDTPLGNRN"

CDS

complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25410))
/gene="CG8465"
/note="CG8465 gene product [alt 3]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAFA8737.1"
/db_xref="GI:7293359"
/translation="MTHQCHRHHDGADSPMLIFTEKSKALDILRHKEARLREFP
NLEOASVVOFGFESIEALKRFCKAKPEKSKPIIISGSGYKSSPTSDNSCSPTGN
GSGFIILPGLSSMSNLISDSTSPSSSSSVNTIANGROOQOQOQOQOQOQOQOQO
GPPRAPPTKOEIYERFKOIEGHIDIKVITINENPFISSGDTPTSLKEGRYAMH
ICAQVANKARIQOILKLTISDREFTOLYVKRSGKACALNLSLDIYINMPDKRG
EIPLHFAKNGHVAWEVLVSPECKSLRNHEKEKEPIICLRNANATHYTIK

alignment_scores:

Quality: 753.50 Length: 323
Ratio: 3.456 Gaps: 4
Percent Similarity: 67.492 Percent Identity: 46.440

alignment_block:

US-09-397-342-1 x AE003506 ..
Align seg 1/1 to: AE003506 from: 1 to: 300994
68 TYAAGlyMetValArgThrAlaLeuGlyIleIleGluGlnGlyPhe 84
|||||.....
295839 TACCGCGGATGGTGGCCACCGCGCTTGCGGATTCGCGGTAGAGAGGCGCC 295888
84 eLeuLysLeuTrpGlnGlyValThrProAlaIleTyArgHisValVal 101
CCTGAACCTGTGGCAGGCGCTTAACGCGCGCTCTACCGACAGCTGCTCT 295938
101 YF..... 101
295939 ATAGTGGCATTCGCCAAGCTCTCTATCTAGCATCTAGATCAGCGCCA 295988
101 101
295989 CAGCAATCTGGCGCAGCAGCTGTAGCATCAGCATAGTACGACAG 296038
102SerglyVal 105
296039 TTCTAACCAACTTCTCCCTGCTCCCATGTCATCCGACGCGGTGTA 296088
105 rGMeValThrTyGluHisLeuArgGluValValPheGlyLysSerGlu 121
296089 GGAATCTGACACTACGACCTGATGCGCAG...GAGTTCACACAGAACGCG 296135
122 AspGluHisTyTrpLeuTrpLysSerValIleGlyIleMetValGlu 138
:::|||||.....
296136 ACCAGAGCTCTGCCGCTTGGAAAGTGCGGCGCTGCGGCGCTCACAGCGCG 296185
138 YValIleGlyGlnPheLeuAlaAsnProThrAspLeuValValGluGln 155
296186 AGCGGTGGCCAGTGGCTTCTGCGCGCGCTGACCTGTAAGGTGCAAA 296235
155 eGluMetGluGlyLysArgLysLeuGluGlyLysProLeuAsnGlyPheArg 171
|||||.....
296236 TCCAAAATGAGGCGCGACGAGCTGTATGAGGCGACCGACGAGGCTGCAC 296285
172 GlyValHisAlaPheAlaLysIleLeuAlaGluGlyIleArg 188
|||||.....
296286 TCGGCTGGCATGCTGCCGCGAGATGTCACGCGTGGCGGAATTAAAGG 296335
188 YLeuTrpAlaGlyTyTrpAlaProAsnIleGlnArgAlaAlaLeuValAsn 205
|||||.....
296336 CCTGTGGAAGGCGACATCCGGAATGTGACGAGCGCGCTGTGTAATC 296385
205 eGlyAspLeuThrTyTrpAspThrValLysHisTyTrpLeuValLeuAsn 221
|||||.....
296386 TGGCGGACCTAACACCATACGACCATCAAGCACTGATCATGATATCGC 296435
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuGlySer 238
:::|||||.....
296436 CTCGACATGCCCGACATCTGTACACAGTGCACGCTGCTGCGCTTGTGCG 296485
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSer 255
|||||.....
296486 CGGATTCGTGGCAGCGATGAGCGACGCGCATGATGTGTGAAGACGC 296535
255 rGlyIleTaNginProArgAspLysGlnGlyArg..... 266
|||||.....
296536 GCATCATGAAACGACCGACGAGAGATGCGCGGTAAAGTACCAATGCG 296585
267Gly... 267
296586 CGCATTTAATCTATTGCAACGCTTAATCATCTATTCTCCGACGAGGCG 296635
268 LeuLeuTyTrpLysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
|||||.....
296636 CTGCTCTATCGGATCCGTGGACTGCTGCGACAGCGGTTTCGAAGG 296685
284 uGlyPheMetSerLeuTyTrpGlyPheLeuProSerTrpLeuArgMet 301

```

296735 296686 GGGCTTTGATGGCCCTGTACAAAGCCTTCTCGCCCTGCTGGATTCAGATG 296735
301 hPrOTtrpsermetvalphettrpleThrTygluylsllaearglumet 317
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296736 CGCGGTGGTCGCGCACCTTCTGTGCTCTGCTTCGACACGATCCGCAAGATG 296785
318 SerGlyValserProphoe 323
|||||:|||||:
296786 ATCGAGACCTCCGCGCTAC 296803
seq_name: -gb_htg4:AC012693
seq_documentation_block:
LOCUS AC012693 79977 bp DNA HTG 03-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC012693
VERSION AC012693.1 GI:6223614
KEYWORDS HTG: HTGS-PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 79977)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDW:10209555 by the submitter.
For further information on this sequence you may e-mail to
fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Source
1..79977
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 22591 a 17930 c 17947 g 21509 t
ORIGIN
Alignment_scores:
Quality: 751.50 Length: 323
Ratio: 3.447 Gaps: 4
Percent Similarity: 67.492 Percent Identity: 46.130
Alignment block:
US-09-397-342-1 x AC012693/rev ..
Align seg 1/1 to reverse of: AC012693 from: 1 to: 79977
68 TTTATGGLMetValArgThrAlaLeuGlyIleIleGluGluGlyPh 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35953 TACCGCGGATGTGGCCACCGCCTTCGGGATTGGCGCGTAGAAGAGGCC 35904
84 eLeuylsLeuTrpGlnGlyValThrProAlaIleTyrrArgHisValValT 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35903 CCTGAAGCTGTGGCAGGCGGTAAACCGCGGCTCTACGACACGTCGTCT 35854
101 YF..... 101
11
35853 ATAGGTGGCATTTCCAAACTCTCATTTCTAGCATCTAGATCACCGCCA 35804
101 ..... 101
35803 CAGGCAATCTGGCCAGCAGCAGCTTAGCATCAGCATTAAGCTCAGCAAG 35754
102 .....SergGlyGlyA 105
|||||:|

```

[illegible]

REFERENCE
AUTHORS

Muscomorpha: Ephydroidea: Drosophilidae: Drosophila.
1 (bases 1 to 260550)

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
ATUES
source

2 (bases 1 to 260550)
Adams,M.D., Celinker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
Brannon,R.C., Rogers,Y.H., Blazer,R.G., Chapple,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
Andrews-Frankoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
Benos,P.V., Bereman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
Botchan,M.R., Bouck,J., Brokstein,P., Brotlier,P., Burtis,K.C.,
Busam,A.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
Pablo,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
Dodson,K., Doup,L.E., Downes,M., Dugan-Kocha,S., Dunkov,B.C.,
Dunp,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S.,
Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J.,
Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
Wel,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
Kernison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
Muzny,D.M., Nelson,D.L., Nelson,K.A., Nixon,K.,
Muzskern,D.R., Pacileb,J.M., Palazolo,M., Plttman,G.S., Pan,S.,
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kimos,I.,
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
Venier,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A.,
Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T.,
Worley,K.C., Wu,D., Yang,S., Yao,O.A., Ye,J., Yeh,R.F.,
Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,C.J.
The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)
20196006

mrna

gene

CDS

1.260550
Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
complement(<174..>467)
/gene="CG13998"
/product="Cr33554"
/db_xref="FLYBASE:FBan0013998"
/db_xref="FLYBASE:FBgn0040949"
/evidence=not_experimental
complement(<174..>467)
/gene="CG13998"
/map="26A8-26A9"
/db_xref="FLYBASE:FBan0013998"
/db_xref="FLYBASE:FBgn0040949"
/evidence=not_experimental
complement(174..467)
/gene="CG13998"
/note="CG13998 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0013998"
/db_xref="FLYBASE:FBgn0040949"
/evidence=not_experimental
/protein_id="AAF52307.1"

mrna

gene

CDS

mrna

gene

CDS

/db_xref="GI:7297038"
/translating="MAKMTFWEFLVALVLTNPMPWRTGAAYTSASMPFLDRHN
GGGHSWSHLPLPTNYSSEMNOQYRRFRQGRMDTPRSKGRQYPPFYAYV"
631..1290
/gene="Ym26ab"
/note="Nucleotide sequence of the *Celera* sequence differs
from the published sequence for this transcript."
/product="Cr25978"
/db_xref="FLYBASE:FBan0009046"
/db_xref="FLYBASE:FBgn0003980"
/gene="Ym26ab"
/note="Ym26ab"
/map="26A8-26A9"
/db_xref="FLYBASE:FBan0009046"
/db_xref="FLYBASE:FBgn0003980"
728..1234
/gene="Ym26ab"
/note="Ym26ab gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript"
/codon_start=1
/db_xref="FLYBASE:FBan0009046"
/db_xref="FLYBASE:FBgn0003980"
/protein_id="AAF52308.1"
/db_xref="GI:7297039"
/translation="MAFNPGHLLINGVALNSVSETIQDPTGQILIPAPLNIIR
SRAATGGAGAPAPATSYSAAPAPAPAPATSYSAAPAPATSYSAAPAPAP
AVSAPAPAPASIPSPPCKNLFLSCPSLPDPSCAPAPASVGSAGAYSQYVPO
AVPEPVEL"
complement(<1708..>2253)
/gene="CG13997"
/product="Cr33553"
/db_xref="FLYBASE:FBan0013997"
/db_xref="FLYBASE:FBgn00031754"
/evidence=not_experimental
complement(<1708..>2253)
/gene="CG13997"
/map="26A8-26A9"
/db_xref="FLYBASE:FBan0013997"
/db_xref="FLYBASE:FBgn00031754"
/evidence=not_experimental
complement(1708..2253)
/gene="CG13997"
/note="CG13997 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0013997"
/db_xref="FLYBASE:FBgn00031754"
/evidence=not_experimental
protein_id="AAF52309.1"
/db_xref="GI:7297040"
/translation="WMKYLGFALLAAGSSJGQAGLEASTGCGTGLRSLPSRSSA
VPSDAYLNRRRQOQIVQEVIVENNFGPGGPGFGPGGPGFGPGGPGFG
RPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGG
KNVPSCEAVIRKVPVPGGYSY"
complement(<3496..>3921)
/gene="Ym26aa"
/product="Cr25988"
/db_xref="FLYBASE:FBan0009048"
/db_xref="FLYBASE:FBgn0003979"
complement(<3496..>3921)
/gene="Ym26aa"
/note="CG9048"
/map="26A8-26A9"
/db_xref="FLYBASE:FBan0009048"
/db_xref="FLYBASE:FBgn0003979"
complement(3496..3921)
/gene="Ym26aa"
/note="Ym26aa gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0009048"
/db_xref="FLYBASE:FBgn0003979"
/evidence=not_experimental
/protein_id="AAF52310.1"

[illegible]

13800 GTACGAGAGTGTCTGGCCACTCGCAATGGGCATCGTCAGAGAGAGGGTCC 13849
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyraRhlValVal 100
:::|||||::: ||||:::|||||::: :::
13850 TGCTGAAGCTCTACGTTGGCATATGCCCATCTGCTTTCCGCGCATCGCTC 13899
101 TyrSerGlyGlyArgMetValThrTygGlnHisLeuArgValValPhe 117
:::|||||::: ::|||:::|||||:::|||||::: :::
13900 TTCACTGGCATCAAATAATGCTGACCTACGACTATATATCGTGAAATAATGAT 13949
117 eglYlysSerGlyAspGlnHisTyrrProLeu.....TrylSerValI 132
:::|||||::: ::|||:::|||||::: :::
13950 CGTGGCCGACAGAGACGGCAGCGCCGACGATATCTCTTCGGGATCTCTGA 13999
132 IeGlyGlyMetLeuAlaGlyValIleGlyGlnPheLeuAlaAsnProThr 148
14000 TCAGTGGCGCTTGTAGCCGGCGCAACTGCCAGCGTACTAACAAGATCCACC 14049
149 AspleuValLysValGlnMetGlnMetGluGlyLysArgLysLeuGlnI 165
:::|||||::: ::|||:::|||||:::|||||::: :::
14050 GAGCTGATTAGAATCCAATATGCAATGAGAGGAGCACGGCGCTCTTAAGGG 14099
165 YlsrProLeuarPheArgrGlyValHisHisAlaPheAlaLysIleLeuA 182
:::|||||::: ||||:::|||||::: ::|||:::|||||::: :::
14100 CGAGCTTCACGCAATTCACATGTGCTCCAGGCCCTCGACTTCATTTACA 14149
182 IaGluGlyGlyIleargGlyLeuTrpAlaGlyTrpValProAsnIleGln 198
14150 GAACGGGGTGGTGGTGGTCTCTGGAAGGAGAACCTGTCCAACACATGG 14199
199 ArgAlaAlaLeuValAsnMet..... 205
:::|||||:::|||||::: ::|||:::|||||::: :::
14200 CCGTCGGCACTCGTAGCCATAGGTGAGCCTTAATAATCGGATTTGATCAG 14249
206GlyAspleut 209
:::|||||:::|||||::: ::|||:::|||||::: :::
14250 TCAGGTCAGATATGTTCAATATCACCCCCTCTCTTAAAGGGAGCGTCA 14299
209 hrThrTyraSprThrValLysHisTyrlLeuValLeuAsnThrProLeuGlu 225
:::|||||::: ||||:::|||||::: ::|||:::|||||::: :::
14300 GCTGCTATGACTTTGGCAACGGCTTCTGATTCGCGAGTTCATCTGGNG 14349
226 AspAsnIleMetThrHisGlyLeuSerSerLeuCysSerglyLeuValAl 242
:::|||||::: ::|||:::|||||::: ::|||:::|||||::: :::
14350 GATAACCCAGAAATGCAATTCGAGCCGCCATACCCGCGCGGTGCGGA 14399
242 aserIleuGlyThrProAlaAspValIleLysSerArgIleMetAsn 259
:::|||||::: ||||:::|||||::: |||||:::|||||::: :::
14400 TGCCTTTTGGACTTCCACAGGACGTGGTCAAGTATCCCGCATCAATGATC 14449
259 InProAtgAspLysGlnGlyIleArgGlyLeuLeuTyrrLysSerSerThrAsp 275
14450 AGCCAACCGAGAGACAGGAGCGGCAATTCATTAACAAGGCGTCCCTGGAT 14499
276 CysLeuIleGlnAlaValGlnGlyGluGlyPheMetSerLeuTyrrLysG 292
:::|||||::: ::|||:::|||||::: |||||:::|||||::: :::
14500 TGCTTAAGTAGTGTGTAGAGAGGAGGCGCTTTTGGCCATGTACAAAGGG 14549
292 yPheLeuProSerThrPheLeuArgmethrProTrpSerMetValPheTrpL 309
:::|||||::: |||||:::|||||::: |||||:::|||||::: :::
14550 ATTAATTCATATGAGTGGTGTGCGACCTGCTCGTGGTGTCTGGA 14599
309 euThrTyrgLysIleargGluMetSergly 319
:::|||||::: |||||:::|||||::: ::|||:::|||||::: :::
14600 TGACCTTCGACAGATTAAGACGGTTTCGTGGC 14631

seq_name: gb_in2:CELK07B1

seq_documentation block:

LOCUS CELK07B1 18557 bp DNA INV 14-MAY-1997

DEFINITION Caenorhabditis elegans cosmid K07B1.

ACCESSION AF003384

VERSION AF003384.1 GI:2088617


```

/evidence-not_experimental
/protein_id="AAB54242.1"
/db_xref="GI:2088824"
/translation="MYGFESSPSSSSDDDFPVYEMKRPDLISLAKKRRALYNOK
ERNRMCVLTSTSLCKHMGESRRKRGRGKRPSRNSVYSTSTDPGVEP
NNMYEPLVQNDQHOCLVAEPIAKKAKTSNFMDFPLEDFPGISRSRYKGNSTKRSLE
DSA"
gene
complement(14557..16723)
/gene="K07B1.7"
CDS
complement(join(14557..14677,14930..15031,15081..15260,
15315..15467,15513..15639,15684..16383,16433..16569,
16624..16723))
/gene="K07B1.7"
/notice="weak similarity to the peptidase family A2"
/codon_start=1
/evidence-not_experimental
/protein_id="AAB54241.1"
/db_xref="GI:2088823"
/translation="MDRTSRNKDILDRFLSKSRNNVEDVSKROEIRCSVAIVNRIP
ELHSDLKRTSRFEVAGQFCFHRHRPEVQAOISGSEKSNITCCCLISLRSEDR
ETLEEFHKKMNSAGVDIPRCFVRLKQSDSEKTSLSLINDIMDIELKPPD
VMPNGNITSPQVLEQILRLPTTVISKLQIOKRRKRGVEVDFYERKDSKAE
KIKCETIIGVDELEDPTEERVNMONPEDNDGDNDQCEMEERHEALHEDYTEOD
RAKPRYEEMEVMEKGGPGLVWYTDKNVDEREKGTCDDMAADMDVNVSYVEG
KSADMDAKAIEMRNDLMSKLEOSAFTKRRPHNLGTRKRRNSDGGSSSEKAS
ENLQKIGGSMIAKMKRGITGLGKNEGRVPAVYVEEDGQSKREKTFGYRGEKL
SRVQKTRHVIASFDVSDREDCEDKQSEELHGEILFRSEVTKMKRFVYLKLO
PPKRNIOKTKTFQNFGLSKCTOLRFLMN"
16923..18437
/gene="K07B1.1"
join(16923..17363,17405..17684,18114..18211,18255..18437)
/gene="K07B1.1"
/notice="similar to the peptidase family S1"
/codon_start=1
/evidence-not_experimental
/protein_id="AAB54236.1"
/db_xref="GI:2088818"
/translation="WILEVDINSNLDDELCSQSTYSFMLTDAAGNTGNPTHLAP
WAVQIRVARKGDEVICGTLITLKHVLAHFOKHFKAKEGSENSKSGRYCES
WVYNSEGLEIYRNORFTDEILRTVYVGMATCPLCEGVMEIRKCVQFE
VYRAIGDFYKTHCEGNDIYILEESTIDVDEANVACPLFEVNIOSGANTSRG
WGSPPGCFDMAAPMIOVLATLETATCEENKTSIPSPSCTAEEEDKNVCSGDS
GGGLTFHOSDSAREFTIAYISGDCYOLIGSSPESRSDINTVRRHOKFIVNFINQA"
BASE COUNT 5630 a 3755 c 3302 g 5870 t
ORIGIN

alignment_scores:
Quality: 627.00 Length: 517
Ratio: 2.498 Gaps: 11
Percent Similarity: 48.549 Percent Identity: 30.561

alignment_block:
US-09-397-342-1 x CELK07B1 ..
Align seg 1/1 to: CELK07B1 from: 1 to: 18557

8 GUUARGLEUleuProleuThrglnArgTyrProArgAlaSer..... 21
:::|||||
4646 CAACAGTCTACTCTTA.....AGAACGATTAATTATTA 4680
22 ...LysPheLeuSerGlyCysAlaIleThrValAlaGluLeuAlaT 37
:::|||||
4681 TTTTCAGTATTTTATCATGTACAGCTCATAGTTGCCGAACGTC 4730
37 hrPheProleuSpleuThrArgLeuGlnMet..... 49
|||||
4731 CATATCTCTGTATTAACAAAACTAGACTACAAATCGCTAGAAATAAG 4780
50 .....GlnGlyGluAlaIleLeuAl 56
:::|||||
4781 TTACAAAGAGAGTATCTGTGATTTTTCGATGATAGTTGTTTTCAT 4830

```

```

56 aArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMetValA 73
|||
4831 TAGTTAAA.....TTTCAGGAGTGTGCT 4855
73 rgrhrAlaLeuGlyIleIleGluGluGluGlyPheLeuLysLeuTyrGln 89
:::|||||
4856 AAGTAACATACATATCATATTAGAGAGAGAACGAAATGGCTTATAGACA 4905
90 GlyValThrProAlaIleTyrArgHisValValTyrSerGlyLysrGme 106
|||||
4906 GAGGTAGCACACACATACACAGACATTTATCTACACAGAAATTCGAT 4955
106 cValThrTyrGluHisLeuArgGluValAlaPheGlyLys..... 119
|||||
4956 GGGTGCTTACAGACAGATTCGTTACTTACATTTAATAAAGTCAGTC 5005
120 .....Ser.GluAspGlnIle 124
5006 ATTTCAATTGTCGAACAATTCAATTTAATTTATTTTCAGCTGAGAAAAGTC 5055
124 sTyrProLeuTyrLysSerValIleGlyGlyMetAlaGlyValIleG 141
|||||
5056 ATTTCCGTTATGGAAGTCTATGCTTTGTGTCGATTTTCTGGCTTATG 5105
141 LysGlnPheLeuAlaAsnProThrAspLeuValLysValGlnMetGlnMet 157
|||||
5106 CTCAAATTCGGCGCTACGTCGACGATTTGTCGAAGCTCAATGCAAAAG 5155
158 GluGlyLysArgLysLeuGluGlyLysProLeuArgPheArgGlyValIle 174
|||||
5156 GAAGGACTTGGAAGATTGCAAGAAACAACCTTTACGATTAACAGGACAAAC 5205
174 shsAlaPheAlaLysIleLeuAlaGluGlyIleArgLysLeuTyrP 191
|||||
5206 CGATTGCTTCAGATCCCTTTATCGAACACAGATTTTTCGATTGTGCA 5255
191 laGlyTyrPylProAsnIleGlnArgAlaAlaLeuValAsnMetGlyasp 207
|||||
5256 TTGGTTGGATGCCAAATTTGACGCGAGCACACTGCTCAATATGCTGAT 5305
208 LeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsn.ThnProL 224
|||||
5306 ATTGTCATATGACAGTGTGAGCATGGCTCATTTGATTAATTTGCAACT 5355
224 eu..... 224
5356 TAAAGTAAGTCTATTGAAAAAAGCGTTTGCACAAAAAATGCGTGT 5405
224 ..... 224
5406 GTCAGAAATGTTTCTTGCGAATTTTATTCGAATTTTACTTTTCAC 5455
224 ..... 224
5456 TTGAATTCAATGAAAAAAGAAATTAATATAGCTAAATCCGCAAGAGA 5505
224 ..... 224
5506 AAAGCAAAAGTTAAAAAATCCAACTTAATAAAAAAAGCGTTTTCACA 5555
224 ..... 224
5556 ACTCGTGTAGTAATTTAAAAAATTTAAAGGAGCCTGAAATTAATCA 5605
224 ..... 224
5606 AAAAAAGTAATAATAATAATACTTTTAAAAAGTAGAATCCTCATTTGTC 5655
224 ..... 224
5656 CGTACAAATGTATAAGTTATGTTGATTTTGAAGAGTACTTTGAAAAAA 5705
224 ..... 224

```

```

5706 AAAATATTTTAAAGTCTCGCGAAATCTATTGTAATAATTTTATATTA 5755
225 .....GluAspAsnIleMethrHisGlyLeuSerS 235
5756 AAACATAACATATCAATTTTACAGGAAATTTGGCTACATCATGCCGTGCTA 5805
225 erLeuCySerGlyLeuValAlaSerIleLeuGlyThrProAlaAspAl 251
5806 GTGCTGTGCTGCAGACTCTGCTGCATTTGATCATCTTCATCAGATGTC 5855
252 IleLysSerArgIleMetAsnGlnProArg..... 261
5856 GTTAAACTCGAATGATGATGATGATGCGACAGAGTGCAGCCCAAAAT 5905
261 ..... 261
5906 GTTGTGTCATCATTTGTTGCTTTCATATTTTCAAAATTTGATTTTCAGG 5955
262 .....AspLysGlnGlyArgGlyLeuLeuTyrlsSerSerThrAs 275
5956 ATGCATAAGAAATAATACCATGTGATCTA...TACAAAGCGCTGCTGCA 6002
275 pCysLeuIleGlnAlaValGlnGlyGlyPheMetSerLeuTyrlsG 292
6003 TTGCTACATAAAATTAATCAAAAGAGATTTTCTCATTACATCAAG 6052
292 lPheLeuProSerThrLeuArgMetThrProThrSerMetValPheTrp 308
6053 GATTCCTCCATCTTATATTCGAATGCGCTGCTCTTACCTTCGCG 6102
309 LeuThrTyrlsGlyLysIleArgGlnMetSerGlyValSerProPhe 323
6103 GTTCCTACGAAGAAATCCGTAATGACCGCGCTTCAAGTTTC 6147
seq_name: gb_prl4:AF155811

seq_documentation_block:
LOCUS AF155811 969 bp mRNA PRI 04-NOV-2000
DEFINITION Homo sapiens mitochondrial uncoupling protein 5 short form mRNA,
complete cds; nuclear gene for mitochondrial product.
ACCESSION AF155811
VERSION AF155811.1 GI:11094338
KEYWORDS
SOURCE
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 969)
Yu, X.X., Mao, W., Zhong, A., Schow, P., Brush, J., Sherwood, S.W.,
Adams, S.H. and Pan, G.
Expression of uncoupling protein homologues, UCP4 and UCP5:
tissue-specific modulation by nutrition and temperature, and
evidence for UCP5 isoforms
Unpublished
JOURNAL
2 (bases 1 to 969)
Pan, G.
Direct Submission
TITLE Submitted (02-JUN-1999) Endocrinology, Genentech, Inc, 1 DNA Way,
JOURNAL Mail Stop-37, South San Francisco, CA 94080, USA
FEATURES
source
1. 969
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq23-q25"
1. 969
/function="reduces mitochondrial membrane potential in
mammalian cells"
/note="UCP5; alternatively spliced"
/codon_start=1
/product="mitochondrial uncoupling protein 5 short form"
/protein_id="AAG29584.1"

```

```

/Db_xref="GI:11094339"
/translacion="MGFPGIIILIFLRKFAATAVIHOKSTVSHMSGLNKPFEVYG
GIASIVAEERETGEPVDLTKRILOYOGOSIDAREIKRGMFALEPRICEGVLAIS
GIAPALROASVYTIKIGTQSKRLFEVLEDETLINNICVSVSSTIANPTD
VKIRMOAGSLTFQSGMIGSFIDYQDEGTRLMRVPAQAAIIVGVELPVYDT
KHLILSGMAGDILTHFVSSFTGCLAGLASNPVVVTRMNOAIVGHVDLYKGT
VDSILKMKKEGFEFALYKGFPMWMLRIGPWNILFETIYBOLKRLQI"
BASE COUNT 258 a 187 c 240 g 284 t
ORIGIN

alignment_scores:
Quality: 557.50 Length: 296
Ratio: 2.617 Gaps: 5
Percent Similarity: 71.959 Percent Identity: 39.865

alignment_block:
US-09-397-342-1 x AF155811 ..
Align seg 1/1 to: AF155811 from: 1 to: 969

23 PheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrPhePr 39
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
121 TTTGATATGCGCGGCGCTTGCCTATCGTGCCTGAGTTGGACCTTCCC 170
39 OleuAspLeuThrLysThrArgLeuGlnMetGlnGlyAlaAlaAlaLeu 56
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
171 TGTGACCTTACCAACACGACCTCAGGTCAGGCCAACCAATGATG 220
56 lAaArgLeuGlyAspGlyAlaArgGlnSerAlaProTyrlsArgGlyMetVal 72
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
221 CCCGTTTC.....AAAGATTAATAATATAGAGGATTTTC 255
73 ArgThrAlaLeuGlyIleIleGluGluGlyPheLeuLysLeuTrpCl 89
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
256 CATGCGCTTTCGCATCTGTAAAGAGAGGCTGATGCGCTCTATTC 305
89 nGlyValThrProAlaIleTyrlsArgHisValAlaValTyrlsGlyArgm 106
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
306 AGCAATGCTCTCGCTGCTGACAGCAACATCATATGCGACCATTAATA 355
106 eValThrTyrlsGlnHisLeuArgGluValAlaPheGlyLysSerGluAsp 122
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
356 TTGGGATTTACCAAGCTTGAAGCGCTTATGCTGAAGCTTTAGAGAT 405
123 GluHisTyrlsProLeuThrLysSerValIleGlyGlyMetMetAlaGly 139
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
406 GAA.....ACTCTTTAATTATGATGATGCTGCGGTAGTGTACAGAGT 449
139 lIleGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
450 GATATCTTCCACTATATGCAATGCCAGCGATGTTGTAAGATTCGAATGC 499
156 lMetGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArgGly 172
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
500 AGCGTCAGAGAGC.....TTGTCGAAGG 525
173 ..ValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgG 188
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
526 AGCATGATTTGGAAGCTTATCATATATACCAACAAGAGCAGCAGCGG 575
188 yLeuThrPheAlaGlyTyrlsAlaProAsnIleGlnArgAlaAlaLeuValAsm 205
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
576 TGTGTGAGAGGGGTGTGCTCAAGCTGCTGCTGCATCGTTGTAG 625
205 eGlyAspLeuThrTyrlsAspThrValLysHisTyrlsLeuValLeuAsn 221
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
626 GAGTAGAGCTACCAAGCTATGATTTACAAAGACATTTATATTTGCA 675
222 ThrProLeuGluAspAsnIleMethrHisGlyLeuSerLeuCySer 238
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
676 GGAATGATGGGAGATACATTTTAACCTGCTTCAGCTTACATG 725

```



```

305 etvalphetrpleuthrtyrgylusilearglumet 317
      ::::|::::|||:::||||::::::::::
932 TCATTTTTTATTACATCAGCACGCTAAAGACTT   969
```


OM of: US-09-397-342-1 to: N_Geneseq_0401.* out_format : pfs
Date: Apr 28, 2001 7:37 PM
About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgml1/USPTO.spool/US09397342/runat.27042001.170001.22942/app_query.fasta_1.384
-DB=N_Geneseq_0401 -QFWT=faststep -SUFFIX=p2n.rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPOP=10.000
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORMEXT -MINLEN=0 -MAXLEN=200000000
-USER=US09397342 -CGML_1199 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPY
-WAIT -THREADS=1

Search information block:

Query: US-09-397-342-1
Query length: 323
Database: N_Geneseq_0401.*
Database sequences: 678276
Database length: 291890651
Search time (sec): 113.230000

score list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SID56/gcgdata/geneseq/NA2000.DAT:A37139		1690.00	3208.05	8.3e-171	998
/SID56/gcgdata/geneseq/NA2000.DAT:A14084		1690.00	3208.63	8.8e-171	1039
/SID56/gcgdata/geneseq/NA2000.DAT:A14085		1674.00	3176.10	5.7e-169	1248
/SID56/gcgdata/geneseq/NA2001.DAT:22721		1186.50	2240.75	7.1e-117	1726
/SID56/gcgdata/geneseq/NA2000.DAT:A46432		668.00	1254.67	6.0e-62	1069
/SID56/gcgdata/geneseq/NA2000.DAT:C39993		668.00	1254.65	6.0e-62	1071
/SID56/gcgdata/geneseq/NA2000.DAT:A45636		638.00	1199.75	1.2e-58	1240
/SID56/gcgdata/geneseq/NA2000.DAT:C40126		638.00	1199.73	1.2e-58	1242
/SID56/gcgdata/geneseq/NA2000.DAT:A37119		560.00	1050.33	1.4e-50	876
/SID56/gcgdata/geneseq/NA2000.DAT:C89945		557.50	1044.85	3.9e-50	936
/SID56/gcgdata/geneseq/NA2000.DAT:C27995		557.50	1044.39	3.9e-50	978
/SID56/gcgdata/geneseq/NA2000.DAT:A27992		557.50	1044.22	3.2e-50	993
/SID56/gcgdata/geneseq/NA2000.DAT:A27996		557.50	1044.05	3.2e-50	1009
/SID56/gcgdata/geneseq/NA2000.DAT:C298032		557.50	1038.68	6.5e-50	1674
/SID56/gcgdata/geneseq/NA2000.DAT:A27999		555.50	1040.00	1.1e-49	1022
/SID56/gcgdata/geneseq/NA2000.DAT:C46816		552.00	1029.73	1.9e-47	1103
/SID56/gcgdata/geneseq/NA2000.DAT:C49402		531.50	992.30	2.5e-47	1124
/SID56/gcgdata/geneseq/NA2000.DAT:C43096		531.50	992.28	2.5e-47	1227
/SID56/gcgdata/geneseq/NA2000.DAT:C63389		515.00	963.37	1.0e-45	957
/SID56/gcgdata/geneseq/NA2000.DAT:C76281		508.50	952.48	4.1e-45	828
/SID56/gcgdata/geneseq/NA1998.DAT:C771710		476.50	891.02	1.1e-41	1220
/SID56/gcgdata/geneseq/NA1999.DAT:C99435		476.50	890.08	1.2e-41	930
/SID56/gcgdata/geneseq/NA1998.DAT:A44595		476.50	884.25	2.6e-41	1612
/SID56/gcgdata/geneseq/NA1999.DAT:C119969		475.50	888.07	1.6e-41	939
/SID56/gcgdata/geneseq/NA1999.DAT:C08507		475.50	887.17	1.8e-41	1172
/SID56/gcgdata/geneseq/NA2000.DAT:C22945		475.50	885.69	2.1e-41	1105
/SID56/gcgdata/geneseq/NA2000.DAT:C246052		475.50	885.53	2.2e-41	1193
/SID56/gcgdata/geneseq/NA2000.DAT:C46055		475.50	885.53	2.2e-41	1193
/SID56/gcgdata/geneseq/NA1998.DAT:C772590		475.50	885.19	2.3e-41	1231
/SID56/gcgdata/geneseq/NA1999.DAT:C84254		475.50	887.38	5.5e-41	2340
/SID56/gcgdata/geneseq/NA1999.DAT:C119968		473.50	884.35	2.6e-41	930
/SID56/gcgdata/geneseq/NA2000.DAT:C50624		473.50	884.35	2.6e-41	930
/SID56/gcgdata/geneseq/NA1999.DAT:C99434		473.50	884.52	3.2e-41	1105
/SID56/gcgdata/geneseq/NA2000.DAT:C229323		473.50	882.52	3.2e-41	1105
/SID56/gcgdata/geneseq/NA1998.DAT:C771712		473.50	881.61	3.6e-41	1204
/SID56/gcgdata/geneseq/NA1999.DAT:C84307		473.50	881.61	3.6e-41	1204
/SID56/gcgdata/geneseq/NA1998.DAT:C09078		472.50	879.25	4.9e-41	1255
/SID56/gcgdata/geneseq/NA1999.DAT:C82381		472.50	876.70	6.8e-41	1596

seq_name: /SID56/gcgdata/geneseq/NA2000.DAT:A37139	ID A37139 standard; cDNA: 998 BP.	AC A37139;	DT 08-AUG-2000 (first entry)	XX	DE Human UCP4 cDNA sequence SEQ ID NO:405.	KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis; transmembrane; secretion; immunoadhesion; pharmaceutical; screening;	XX	OS Homo sapiens.	XX	PN W0200012708-A2.	PD 09-MAR-2000.	XX	PF 01-SEP-1999; 99W0-US20111.	XX	01-SEP-1998; 98US-0098716.	PR 01-SEP-1998; 98US-0098749.	PR 01-SEP-1998; 98US-0098750.	PR 02-SEP-1998; 98US-0098803.	PR 02-SEP-1998; 98US-0098821.	PR 02-SEP-1998; 98US-0098843.	PR 09-SEP-1998; 98US-0099536.	PR 09-SEP-1998; 98US-0099596.	PR 09-SEP-1998; 98US-0099598.	PR 09-SEP-1998; 98US-0099602.	PR 09-SEP-1998; 98US-0099642.	PR 10-SEP-1998; 98US-0099741.	PR 10-SEP-1998; 98US-0099754.	PR 10-SEP-1998; 98US-0099763.	PR 10-SEP-1998; 98US-0099792.	PR 10-SEP-1998; 98US-0099808.	PR 10-SEP-1998; 98US-0099812.	PR 10-SEP-1998; 98US-0099815.	PR 15-SEP-1998; 98US-0100385.	PR 15-SEP-1998; 98US-0100388.	PR 15-SEP-1998; 98US-0100380.	PR 16-SEP-1998; 98US-0100384.	PR 16-SEP-1998; 98US-0100627.	PR 16-SEP-1998; 98US-0100661.	PR 16-SEP-1998; 98US-0100662.	PR 16-SEP-1998; 98US-0100664.	PR 17-SEP-1998; 98US-0100683.	PR 17-SEP-1998; 98US-0100684.	PR 17-SEP-1998; 98US-0100710.	PR 17-SEP-1998; 98US-0100711.	PR 17-SEP-1998; 98US-0100919.	PR 17-SEP-1998; 98US-0100930.	PR 18-SEP-1998; 98US-0100848.	PR 18-SEP-1998; 98US-0101014.	PR 18-SEP-1998; 98US-0101068.	PR 22-SEP-1998; 98US-0101071.	PR 22-SEP-1998; 98US-0101279.	PR 23-SEP-1998; 98US-0101471.	PR 23-SEP-1998; 98US-0101472.	PR 23-SEP-1998; 98US-0101474.	PR 23-SEP-1998; 98US-0101475.	PR 23-SEP-1998; 98US-0101476.
--	-----------------------------------	------------	------------------------------	----	--	--	----	------------------	----	--------------------	-----------------	----	-------------------------------	----	----------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------

134 IymetMetAlaGlyValIleGlyInPheLeuAlaAsnProThrAspLeu 150
 427 GGATGATGCGTGGTGTATTGGCCAGTTTTCACCAATCCAGTACCTA 476
 151 ValIysValGlnMetGlnMetGluGlyLysArgLysLeuGluGlyLysPr 167
 477 GTGAAGGTTCAGATGCAAAATGGAAGAAAAAGAACTGGAAGAAAAACC 526
 167 OleuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
 527 ATTCGCAATTCGTGTGATCATGATTCGCAAAAATCTTACCTGAG 576
 184 LysGlyIleArgGlyLeuThrPheAlaGlyThrPheAlaProAsnIleGluArgAla 200
 577 GAGGAAATACGAGGGCTTTGGGAGGCTGGTACCAATATACAAAGACCA 626
 201 AlaLeuValAsnMetGlyAspLeuThrThrTyAspThrValIysHisLys 217
 627 GCACCTGGTGAATATGGAGATTAAACCATATGATACAGTGAACACTA 676
 217 rLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisGlyLeu 234
 677 CTGTGTATGTAATACACCACTTGAGGACAAATATCATGACTCAGCGTTAT 726
 234 eSerLeuCysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
 727 CAAGTTATGTTCTGAGCTGTAGCTCTATCTGAGGACACCGCAT 776
 251 ValIleYsSerArgIleMetAsnGlnProArgAspLysGlnGlyArgI 267
 777 GTCAATCAAAAGCAGAAATATGAAATCAACACGAGATTAACAGGAAGGG 826
 267 YleuLeuTyIysSerSerThrAspCysLeuIleGlnAlaValGlnGly 284
 827 ACTTTGTAATTAATCATCGACTGCTGATTCAGCTCTCAAGGTG 876
 284 LuGlyPheMetSerLeuTyIysGlyPheLeuProSerThrLeuArgMet 300
 877 AAGATTCATCAGTCTATATTAAGGCTTTTACCATCTTGCGTGAATG 926
 301 ThrProTrpSerMetValPheTrpLeuThrTyrgLysIleArgGlu 317
 927 ACCCTTGCTCAATGCTGTCTGCTTACTTATGAAAAAATCAGAGAGAT 976
 317 tSerGlyValSerProPhe 323
 977 GAGTGGAGTCACTCATTT 995

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT.A14084

seq_documentation_block:

ID A14084 standard; CDNA; 1039 BP.

AC A14084;

DT 21-JUL-2000 (first entry)

DE CDNA encoding human uncoupling protein UCP4.

XX Uncoupling protein; UCP4; expressed sequence tag; EST; human;

KW mitochondrial membrane; proton leakage; heat production; metabolic rate;

KM drug screening; obesity; stroke; trauma; sepsis; infection; ss.

XX Homo sapiens.

XX Key location/Qualifiers

FT CDS 40..1011

FT /tag= a

PN /product= "Human UCP4"

WO200017353-A1.

XX 30-MAR-2000.
 PD 99MO-0521194.
 XX
 PF 15-SEP-1999; 99MO-0521194.
 XX
 PR 22-SEP-1998; 98US-0101279.
 PR 30-DEC-1998; 98US-0114223.
 PR 16-APR-1999; 99US-0129674.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Adams S, Pan J, Zhong A;
 DR WPI: 2000-292842/25.
 DR P-PDB; Y81497.
 XX
 PT New nucleic acid encoding human uncoupled protein-4, useful e.g for
 PT identifying metabolic regulators for treatment of obesity
 PS Claim 3; Fig 2; 80pp; English.

XX This sequence represents cDNA encoding human uncoupling protein UCP4. The
 CC human UCP4 cDNA (ATCC 203134) was isolated from a brain cDNA library
 CC using a probe generated using primers A14086-A14087. These primers were
 CC based on a UCP4 "from DNA" sequence (A14085) derived from a number of
 CC ESTs (expressed sequence tags) which were selected on the basis of
 CC homology with human UCP3. The human UCP4 gene has been mapped to
 CC chromosome 6p11.2-q12. UCP4 catalyses the leakage of protons through
 CC the mitochondrial membrane, thus bypassing ATP synthase and thereby
 CC reducing the efficiency of ATP synthesis. Modulation of UCP4 activity or
 CC via modulation of ATP synthetic efficiency. UCP4 nucleic acids may be
 CC used for recombinant production of UCP4 and as a source of primers and
 CC hybridisation probes which may be used for the analysis of UCP4
 CC expression, for screening for homologous sequences, and for chromosome
 CC or gene mapping. They can be also be used to produce transgenic or
 CC knockout animals for the development and screening of therapeutic agents,
 CC as a source of antisense nucleotides, and in gene therapy for metabolic
 CC disorders. The UCP4 protein can be used to screen for specific modulators
 CC of activity and to raise antibodies which may be used therapeutically as
 CC UCP4 antagonists or UCP4 targeting proteins, for affinity purification
 CC of UCP4, and as immunoassay reagents for detecting UCP4 expression,
 CC e.g., for the diagnosis of impaired neural activity or neural
 CC degeneration. Agents that modulate UCP4 activity are used to control the
 CC metabolic rate in mammals. UCP4 upregulators can be used to increase the
 CC metabolic rate in order to treat obesity and the symptoms associated
 CC with stroke, trauma, sepsis and infection.

XX Sequence 1039 BP; 294 A; 215 C; 271 G; 259 T; 0 other;

alignment_scores:

Quality: 1690.00 Length: 323

Ratio: 5.232 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-397-342-1 x A14084 ..

Align seg 1/1 to: A14084 from: 1 to: 1039

1 MetSerValProGluGluGluArgLeuLeuProLeuThrGlnArgTr 17
 40 ATGTCTCCGTCGGAGGAGGAGGAGGCTTTTGGCGCTGACCCAGATG 89
 17 ProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
 90 GCCCGGAGGAGGAGGAGGAGGAGGCTTTTGGCGCTGACCCAGATG 139
 34 LuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
 140 AGCTAGCAACCTTTCCCTGTGATCTCACAAAAAATCGACTCCAAATGCA 189

```

51 GlyAlaAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgLysSerAlaPr 67
190 GGAGAACACCTCTTCTCGTTGGAGAGCGTGCAGAGAAATCTGCCCC 239
67 CTATAGGAGATGCTGGCAGACAGCCCTAGGATCATTTGAAGAGAGAGCT 289
240 CTATAGGAGATGCTGGCAGACAGCCCTAGGATCATTTGAAGAGAGAGCT 289
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValVal 100
290 TTCTAAAGCTTTGGCAAGGAGTACACCCCATTTTACAGACAGCTAGTG 339
101 TyrSerGlyLysArgMetValThrTyrGlnHisLeuArgLysValAla 117
340 TATTTGGAGGTGGAATGTCACATATGAAATCTCCGAGAGGTGTGTT 389
117 eGlyLysSerGlyAspGlyHisTyrProLeuTyrLysSerValIleGly 134
390 TGGCAAAAGTGAAGATGACATATTCCTTTGGAATCATGCTATTGGAG 439
134 LysMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
440 GGATGATGCTGCTGTTATGGCCAGTTTTCAGCAATCCAGTACGACCTA 489
151 ValLysValGlnMetGlnMetGlnGlyLysArgLysLeuGlnGlyLys 167
490 GTGAGAGTTCAGATGCAATGGAAGGAAAGAAACCTGAGAGGAAACC 539
167 GluPheArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAla 184
540 ATTGCCATTTCTGCTGCTACATCATGATTTGCAAAAATTTAGCTGAG 589
184 LysIleArgGlyLeuTyrPalaGlyTyrValProAsnIleGlnArgAla 200
590 GAGGAATACAGAGGCTTTGGCAGCGCTGGTACCACATATACAAAGACA 639
201 AlaLeuValAsnMetGlyAspLeuThrTyrAspThrValLysHisTyr 217
640 GCACCTGATGATGAGATGGAATTTACCACTATGATACACTGAAACCTA 689
217 IleValIleAsnThrProLeuGlnAspAsnIleMetThrHisGlyLeu 234
690 CTGGATTCATATACACACTTACGACCAATATCATGACTACGAGGTTAT 739
234 eTyrSerLysSerGlyLeuValAlaSerIleLeuGlnTyrProAlaAsp 250
740 CAAGTTATGCTCTGACCTGAGCTTCTATCTCGGAAACACAGCCGAT 789
251 ValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArg 267
790 GTGATCAAAAGCAGAAATTAAGATCAACCAAGATTAACAAAGAGAGGG 839
267 LysLeuLysSerThrAspCysLeuIleGlnAlaValGlnGlyG 284
840 ACTTTGTATTAATCATCGACTACGCTTGTGATTCAGGCTTCAAGGTG 889
284 LysIlePheMetSerLeuTyrLysGlyPheLeuProSerTyrLeuArgMet 300
890 AAGGATTCATGAGCTATATTAAGGCTTTTACCATCTTGCGTGAAGATG 939
301 ThrProTyrSerMetValPheTyrPheLeuThrTyrGlnLysIleArg 317
940 ACCCTTGTGCAATGCTGCTGCTTACTTATGAAGAAATCAGAGAGAT 989
317 tSerGlyValSerProPhe 323
990 GAGTGGAGTCAGTCCATTT 1008

```

seq_name: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT: A14085

seq_documentation_block:

ID A14085 standard; DNA: 1248 BP.
XX

```

AC A14085;
XX
XX 21-JUL-2000 (first entry)
DT
DE Human uncoupling protein UCP4 "from DNA" sequence, SEQ ID NO:5.
XX
XX
XX Uncoupling protein; UCP4; expressed sequence tag; EST; UCP3 homologue;
KW human; chromosome 6p11.2-q12; ATP synthesis; energy efficiency;
KW mitochondrial membrane; proton leakage; heat production; metabolic rate;
KW drug screening; obesity; stroke; trauma; sepsis; infection; ds.
XX
XX Homo sapiens.
OS
XX WO200017353-A1.
XX
XX 30-MAR-2000.
XX
XX 15-SEP-1999; 99WO-US21194.
XX
XX 22-SEP-1998; 98US-0101279.
XX 30-DEC-1998; 98US-0114223.
XX 16-APR-1999; 99US-0129674.
XX
XX (GERTH ) GENENTECH INC.
XX
XX Adams S, Pan J, Zhong A;
XX
XX WPI: 2000-292842/25.
XX
XX New nucleic acid encoding human uncoupled protein-4, useful e.g for
XX identifying metabolic regulators for treatment of obesity.
XX
XX Example 1: Fig 7: 80pp; English.
XX
XX The invention relates to human uncoupling protein UCP4 (Y81497) and
XX cDNA encoding it (A14085). The human UCP4 cDNA (ATCC 203134) was isolated
XX from a brain cDNA library using a probe generated using primers
XX A14086-A14087. These primers were based on a UCP4 "from DNA" sequence
XX (A14085) derived from a number of ESTs (expressed sequence tags) which
XX were selected on the basis of homology with human UCP3. The human UCP4
XX gene has been mapped to chromosome 6p11.2-q12. UCP4 catalyses the leakage
XX of protons through the mitochondrial membrane, thus bypassing ATP
XX synthase and thereby reducing the efficiency of ATP synthesis. Modulation
XX of UCP4 activity or expression can therefore alter the metabolic rate and
XX heat production via modulation of ATP synthetic efficiency. UCP4 nucleic
XX acids may be used for recombinant production of UCP4 and as a source of
XX primers and hybridisation probes which may be used for the analysis of
XX UCP4 expression, for screening for homologous sequences, and for
XX chromosome or gene mapping. They can be also be used to produce
XX transgenic or knockout animals for the development and screening of
XX therapeutic agents, as a source of antisense nucleotides, and in gene
XX therapy for metabolic disorders. The UCP4 protein can be used to screen
XX for specific modulators of activity and to raise antibodies which may be
XX used therapeutically as UCP4 antagonists or UCP4 targeting proteins, for
XX affinity purification of UCP4, and as immunosay reagents for detecting
XX UCP4 expression, e.g., for the diagnosis of impaired neural activity or
XX neural degeneration. Agents that modulate UCP4 activity are used to
XX control the metabolic rate in mammals. UCP4 upregulators can be used to
XX increase the metabolic rate in order to treat obesity and the symptoms
XX associated with stroke, trauma, sepsis and infection. The present
XX sequence represents the human UCP4 "from DNA" sequence assembled from EST
XX sequences with homology to human UCP3.
XX
XX Sequence 1248 BP: 357 A; 263 C; 301 G; 326 T; 1 other.

```

alignment_scores:

Quality: 1674.00 length: 323
Ratio: 5.215 Gaps: 0
Percent Similarity: 99.381 Percent Identity: 99.381

alignment_block:

US-09-397-342-1 x A14085

Align seg 1/1 to: A14085 from: 1 to: 1248

```

1 MetSerValProGluGluGluArgLeuLeuProLeuThrGlnArgTyr 17
  |||
26 ATGTCCTCCCGGAGGAGAGAGAGAGCTTTCGCGCTGACCCAGAGATG 75
17 ProArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaG 34
  |||
76 GCCCGGAGGAGCAATTCCTACTGCTCGGCTGCGGCTACCGCTGCGG 125
34 IuLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGln 50
  |||
126 AGCTAGACAACTTCCCTGGATCTCACAACAACTGCAGCTCCAAATGCA 175
51 GlyGluAlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGlnSerAlaPr 67
  |||
176 GGAGAAACACCTTCTGCTCGTGGAGAGCGGTGCAAGAGAACTGCCCC 225
67 OTYrArgGlyMetValArgThrAlaThrAlaGlyIleIleGluGluGluGlyP 84
  |||
226 CTATAGGGGAGATGTCGCCACAGCCCTAGGATCATTTGAGAGAGAGGCT 275
84 heLeuLysLeuTrpGlnGlyValThrProAlaIleTyrArgHisValVal 100
  |||
276 TTCTAAAGCTTTGGCAGAGAGTGACACCCGCCATTTCAGACAGCTAGTT 325
101 TyrSerGlyGlyArgMetValThrTyrGluHisLeuArgGluValValPh 117
  |||
326 ATTCTGTGAGGTGCAATGTCACATATGAAACATCTCCGAGAGTGTGTT 375
117 eGlyLysSerGluAspGlnHisTyrProLeuTrpLysSerValIleGlyG 134
  |||
376 TGGCAAAAGTGAAAGATGACATTATCCCTTGGAAATCAGTCATTGGAG 425
134 LymMetAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAspLeu 150
  |||
426 GGATGATGGCGGTGTTATGGCCAGTTTTCAGCAATCCACAGACCTA 475
151 ValLysValGlnMetGlnMetGlyLysArgLysLeuGluGlyLysPr 167
  |||
476 GTGAGAGTTCAGATGCAAAATGGAGAAAGAAAGAACTGGAGAGAAACC 525
167 OLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAlaGlu 184
  |||
526 ATTTCGATTTCGTGCTGCTACATCATGCAATTTGCCAAAATCTTACCTG 575
184 LysGlyIleArgGlyLeuTrpAlaGlyTyrValProAsnIleGlnArgAla 200
  |||
576 GAGCAATACGAAAGCTTTGGGCAAGCTGGGTACCCCAATATACAAAGCA 625
201 AlaLeuValAsnMetGlyAspLeuThrThrTyrAspThrValLysHisTyr 217
  |||
626 GCATCGGTGATATGAGAGATTTAACCACTTATATACAGGAAACACTA 675
217 rLeuValLeuAsnThrProLeuGlnAspAsnIleMetThrHisGlyLeuS 234
  |||
676 CTTCGTTCATACACCACTTGAGGACAAATATATGATGACTCAGCGTTAT 725
234 eSerLeuCysSerGlyLeuValAlaSerIleLeuGlyThrProAlaAsp 250
  |||
726 CAAGTTTATGTTGGAAGTGTGCTTCTATTTGGGAGAACCCAGCGCAT 775
251 ValLysSerArgIleMetAsnGlnProArgAspLysGlnGlyArgGly 267
  |||
776 GTCTCAAAAGCAGAAATATGAAATCAACACAGATTAACCAAGAGAGCG 825
267 YLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValGlnGlyG 284
  |||
826 ACTTTGTATTAATCATCGACTGCTGATTCAGGCTGTTCAGAGGTG 875
284 LngLysPheMetSerLeuTyrLysGlyPheLeuProSerTrpLeuArgMet 300
  |||

```

```

876 AAGCATTCATGAGTCTATATATAAGCTTTTACCATCTTGCGTGAGATG 925
301 ThrProTrpSerMetValPheTrpLeuThrTyrGluLysIleArgGlu 317
  |||
926 ACCCTTGGTCAATGCTGTCTGCTTACTTATGAAAAATCAGAGAGAT 975
317 tSerGlyValSerProPhe 323
  |||
976 GAGTGGAGTCACTTCATT 994

```

seq_name: /SID6/gcgdata/geneseq/geneseqn/NA2001.DAT.F27721

seq_documentation_block:

ID F27721 standard; cDNA: 1726 BP.

XX F27721;

DT 28-MAR-2001 (first entry)

DE Human transport protein TPPT-21 coding sequence.

XX Human; transport protein; TPPT; transport disorder; metabolic disorder;

KW neurological disorder; cardiovascular disorder; reproductive disorder;

XX immune disorder; cancer; ss.

OS Homo sapiens.

XX WO200078953-A2.

PD 28-DEC-2000.

XX 16-JUN-2000; 2000MC-US16668.

PF 17-JUN-1999; 990S-0139923.

PR 10-AUG-1999; 990S-0148177.

PR 18-AUG-1999; 990S-0149357.

PR 28-OCT-1999; 990S-0162287.

XX (INCY-) INCYTE GENOMICS INC.

XX Lal P, Yang J, Yue H, Hillman JL, Tang YR, Bandman O, Burford N;

PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;

XX WPI: 2001-041424/05.

DR P-PSDB; B60101.

XX Isolated polypeptide with a human transport protein sequence is useful

PT for the diagnosis, prevention and treatment of disorders associated

XX with the immune, reproductive and cardiovascular systems -

PS Claim 5; Page 153-154; 165pp; English.

XX The present invention provides the protein and coding sequences for 43

CC novel human transport proteins (designated TPPTs). These can be used in

CC the diagnosis and treatment of transport, metabolic, neurological,

CC reproductive, cardiovascular and immune disorders, and cell proliferative

XX disorders such as cancer.

SQ Sequence 1726 BP; 552 A; 302 C; 399 G; 473 T; 0 other;

alignment_scores:

Quality: 1186.50

Length: 244

Ratio: 5.181

Gaps: 1

Percent Similarity: 93.852

Percent Identity: 92.623

alignment_block:

US-09-397-342-1 x F27721 ..

Align seg 1/1 to: F27721 from: 1 to: 1726

95 ILeTyrArgHisVal..... 99

397 ATACATACATATACATATACATATACCAAGTACCTTTTATATATCT 446
100ValTyrSerGlyIyArGmetValIhrrTyrgLuhIsLeua 113
447 TTTTACCAGGTGTATTCGAGGTGCAATGTCACATATGAACTCTCC 496
113 rgluValValIhneGlyIySerGluAspGluHisTyrProLeuTyrLys 129
497 GAGAGCTGTGTGGCAAAAGTGAAGATGACATTATCCCTTTGGAAA 546
130 SerValIleGlyIyMetMetIaGlyValIleGlyIyPheLeuIaAs 146
547 TCAGTATTGGAGGATGATGCTGTGTATTGGCCAGTTTACCCTAA 596
146 nProThrAspLeuValIyValGlnMetGlnMetGluGlyIySarGlyLys 163
597 TCCAATGACCTAGTATGAGTTTCAGATGCAATGCAAGCAAAAGCAAA 646
163 euGluGlyIyProLeuArgPheArgGlyValIhHisIaPheAlaLys 179
647 TCGAAGCAAAACCATTCGATTTCGTGTACATCATCATTCGCAAAA 696
180 IleLeuAlaGluGlyIyIleArgGlyLeuTrrPalaGlyTrrValProAs 196
697 ATCTTAGCTGAAGAGGAATACAGGGCTTTGGCAGGCTGGGTACCCAA 746
196 nIleGlnArGAlaIaLeuValaAsnMetGlyAspLeuThrTrrTyrAspT 213
747 TATACAAAGAGCAGCAGCTGTGAATATGGAGATTAAACCATTATGATA 796
213 hValIyHisTyrLeuValIleuAsnThrProLeuGluAspAsnIleMet 229
797 CAGTGAACACTACTGTGATTGAATACCACTTGAGGACATATCATG 846
230 ThrHisGlyLeuSerSerLeuGlySerGlyLeuValaIaSerIleuGlu 246
847 ACTCAGGCTTATCAAGTTATGTTCTGACAGCTGTGCTTCTTATTCGG 896
246 yThrProIaAspValIleLysSerArgIleMetAsnGlnProArgAspL 263
897 AACACAGCCGATGTCATCAAAAGCAGATTAATGATCAACGAGATA 946
263 yGlnGlyIyArgGlyLeuLeuTyrIySerSerThrAspGlyLeuIleGln 279
947 AACCAAGAGGGGACTTTGTATTAATCATCAGCTGACTGTGATTGAG 996
280 AlaValGlnGlyIyGluIyPheMetSerLeuTyrIyGlyIyPheLeuProse 296
997 GCTGTTCAGAGTGAGGATTTCAGTCTATATAAAGGCTTTTACCATC 1046
296 rTrrPleuArgMetThrProTrrSerMetValPheTrrPleuThrTyrGluL 313
1047 TTGGCTGGAATGACCCCTTGTGCAATGCTGTCTGCTTACTTATGAAA 1096
313 ySllLeArgLumetSerGlyValaSerProPhe 323
1097 AATTCAGAGATGATGAGTCAAGTCCATT 1128
seq_name: /Sids6/gcdata/geneseq/geneseqn/NA2000.DAT:C46432
seq_documentation_block:
ID C46432 standard; DNA: 1069 BP.
XX C46432;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 50118.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX

OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157785.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158322.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores: Quality: 668.00 Length: 333
 Ratio: 2.760 Gaps: 9
Percent Similarity: 72.673 Percent Identity: 42.342

alignment_block:
US-09-397-342-1 x C39993 ..

Align seg 1/1 to: C39993 from: 1 to: 1071

2 SerValProGluGluGluGluArgLeuLeuProLeuThrGlnArgTyrPr 18
 ||||| ||||| :||||| ||
17 TCAAATCCGGTCAAAATGAGCGG...AGCCGAGTGCACGAGAACCC 63
18 cArGAlaserLysPheLeuLeuSerGlyCysAlaIatThrValaIaGlu 35
 | :||||| :||||| |||||
64 GACCGAGAACGAGATCGTCTGCTCACTCTCTCGATGCTGACAGAT 113
35 euAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMeGlnGly 51
 :||||| :||||| :||||| ||
114 CAGTACGTTTCCGATGATCTCACAAAGACTAGATCAGTCCATCGT 163
52 GuAlaAlaLeuAla.....ArgLeuGlyAspGlyAlaArgGluSe 65
 :||||| :||||| :||||| ||
164 TCGGATCCGCTTCTGCTGCGCATCGATGCGC..... 196
65 rAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81
 ||||| :||||| :||||| ||

```

197 .....GCTTCGAGTCGTATCGAGA 218
82 .....GlutylPheLeuYsLeuTrpGlnGlyValThProAla 94
219 TTGGCAAGAGAGAGAGATGTGTCTCTCAAGGCTATCTCCGCG 268
95 ILeTyrAlaGhisValValTyrSerGlyGlyArgMetValThTyrGluH 111
269 ATATACAGACACCTGTCTACACGCTATCAGAAATCATGTGATACGAGA 318
111 sLeuArgGluValValPheGlyYsSerGlu.....AspGluHisT 125
319 TTGAAAGACTCATGCTC...AGATCTGAACCTAACATACGACGCTC 365
125 yrProLeu...TrpYsSerValIleGlyGlyMetMetAlaGlyValIle 140
366 TTCTCTCGGCACAAAGCTCTCTCGAGAGA...TTTCTGGGTATATA 412
141 GLYGlnPheLeuAlaAsnProThrAspLeuValYsValGlnMetGlnMe 157
413 GCTCAGCTAGTGGCTAGTCCAGCTGATTTGGTCAAGTGAATGCAAGC 462
157 tGluGlyAlaArgYsLeuGlyYsProLeuArgPheArgGlyValH 174
463 AGATGGTAGATTGGTAGCCCAAGGCTGAACCGAGTACTCGGACCAA 512
174 tshAlaPheAlaYsIleLeuAlaGlnGlyIleArgGlyLeuTrp 190
513 TCGAGGCTTTTACCAAAATCTCTCAATCAGAGAGATAAAGGTTATGG 562
191 AlAGlyTrpValProAsnIleGlnAlaAlaLeuValAsnMetGlyAs 207
563 AAGGCTTCTTCCAAACATCCAGAGCATTCTAGTGAATGAGAGA 612
207 PleuThrThrTyrAspThrValYsHisTyrLeuValLeuAsnThrPro 224
613 ACTAGCTGCTATGATCAGCCCAACACTTGTGATCATTAGAGAAATGG 662
224 euGluAspAsnIleMetThrHisGlyLeuSerSerLeuCySerGlyLeu 240
663 CTGAGGATACATTTTGGCAGACCTTGTCTTATATATGTCGGCTT 712
241 ValAlaSerIleLeuGlyThrProAlaAspValIleYsSerArgIleMe 257
713 GCCTCGCAAGTTGAGTGTCCAGCTGATGTGTGAAGACGAGATGAT 762
257 tAsnGlnProArgAspYsGlnGlyArgGlyLeuLeuTyrYsSerSerT 274
763 GAAC.....CAGGGTGAATAATGCTGTGTACAGAAATTTCTT 797
274 hrAspCysLeuIleGlnAlaValAlaGlnGlyGlnGlyPheMetSerLeuTr 290
798 ACAGACTTGTGTGAAGACGCTTAACTTGAAGAAATAAGCTTTGTG 847
291 LyGlyPheLeuProSerTrpLeuArgMetThrProTrpSerMetValPh 307
848 AAAGGTTTCTTCCGACATGGGCAAGGCTTGACCGTCCGATTCGAT 877
307 eTrpLeuThrTyrGlnYsIleArgGlnMetSerGlyValSerProPhe 323
898 TTGGGCTCTCTATGAGAGATTAGACTGCTGGCAGGAGATCTCTCCCTC 946
seq_name: /SID6/gcdata/geneseq/geneseqn/NA2000.DAT:CA5636
seq_documentation_block:
ID CA5636 standard; DNA: 1240 BP.
XX
XX CA5636;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 47214.
XX

```

```

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134321.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.

```

PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	29-JUN-1999;	9905-0140991
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142390
PR	08-JUL-1999;	9905-0142803
PR	12-JUL-1999;	9905-0142920
PR	19-JUL-1999;	9905-0142977
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334
PR	21-JUL-1999;	9905-0144335
PR	21-JUL-1999;	9905-0144352
PR	22-JUL-1999;	9905-0144632
PR	22-JUL-1999;	9905-0144632
PR	22-JUL-1999;	9905-0144864
PR	22-JUL-1999;	9905-0144814
PR	23-JUL-1999;	9905-0145192
PR	23-JUL-1999;	9905-0145145
PR	23-JUL-1999;	9905-0145086
PR	23-JUL-1999;	9905-0145087
PR	26-JUL-1999;	9905-0145224
PR	26-JUL-1999;	9905-0145276
PR	27-JUL-1999;	9905-0145913
PR	27-JUL-1999;	9905-0145918
PR	27-JUL-1999;	9905-0145913
PR	28-JUL-1999;	9905-0145919
PR	28-JUL-1999;	9905-0145951
PR	02-AUG-1999;	9905-0145386
PR	02-AUG-1999;	9905-0145388
PR	02-AUG-1999;	9905-0145389
PR	03-AUG-1999;	9905-0145638
PR	03-AUG-1999;	9905-0147004
PR	04-AUG-1999;	9905-0147204
PR	04-AUG-1999;	9905-0147302
PR	05-AUG-1999;	9905-0147192
PR	05-AUG-1999;	9905-0147260
PR	06-AUG-1999;	9905-0147303
PR	06-AUG-1999;	9905-0147416
PR	09-AUG-1999;	9905-0147493
PR	09-AUG-1999;	9905-0147935
PR	10-AUG-1999;	9905-0148171
PR	11-AUG-1999;	9905-0148319
PR	12-AUG-1999;	9905-0148341
PR	13-AUG-1999;	9905-0148565
PR	13-AUG-1999;	9905-0148684
PR	16-AUG-1999;	9905-0149368
PR	17-AUG-1999;	9905-0149175
PR	18-AUG-1999;	9905-0149426
PR	20-AUG-1999;	9905-0149722
PR	20-AUG-1999;	9905-0149723
PR	20-AUG-1999;	9905-0149929
PR	23-AUG-1999;	9905-0149902
PR	23-AUG-1999;	9905-0149930
PR	25-AUG-1999;	9905-0150564
PR	26-AUG-1999;	9905-0150584
PR	27-AUG-1999;	9905-0151065
PR	27-AUG-1999;	9905-0151066
PR	27-AUG-1999;	9905-0151080
PR	31-AUG-1999;	9905-0151308
PR	31-AUG-1999;	9905-0151438

PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155133;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155659;
PR	28-SEP-1999;	9905-0156596;
PR	29-SEP-1999;	9905-0157117;
PR	04-OCT-1999;	9905-0157173;
PR	05-OCT-1999;	9905-0157865;
PR	06-OCT-1999;	9905-0158029;
PR	07-OCT-1999;	9905-0158232;
PR	08-OCT-1999;	9905-0158369;
PR	12-OCT-1999;	9905-0158293;
PR	13-OCT-1999;	9905-0159294;
PR	14-OCT-1999;	9905-0159295;
PR	13-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159530;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160747;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160981;
PR	22-OCT-1999;	9905-0160989;
PR	25-OCT-1999;	9905-0161404;
PR	25-OCT-1999;	9905-0161405;
PR	25-OCT-1999;	9905-0161406;
PR	26-OCT-1999;	9905-0161359;
PR	26-OCT-1999;	9905-0161360;
PR	26-OCT-1999;	9905-0161361;
PR	28-OCT-1999;	9905-0161920;
PR	28-OCT-1999;	9905-0161992;
PR	29-OCT-1999;	9905-0161993;
PR	29-OCT-1999;	9905-0162142;

```
alignment_scores:      Length:      336
      Quality:      638.00
      Ratio:      2.626
      Gaps:      11
Percent Similarity:      72.321
Percent Identity:      41.964
```

alignment_block:

Align seg 1/1 to: C45636 from: 1 to: 1240

[illegible]

```

65 rAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81
235 .....GCTTCGGAGTGTATCGGAGA 256
82 .....GluGlyPheLeuIleuTPrpGlnGlyValThrProAla 94
257 TTGGAGAAAGAGAGAGTGTGTCTACAAAGGTCTATCTCCGGCG 306
95 IleTyrArgHisValValTyrSerGlyIleArgMetValThrTyrGluH 111
307 ATTATCAGACACCTGTCTACAGCCATATCATGTGATACGAGAA 356
111 sIeuArgGluValValPheGlyLysSerGlu.....AspGlnHisT 125
357 TTGAAAGGACATCGTC...AGATCTGAACCTAACATATGCGAGCTC 403
125 YrProLeu...TPrpLysSerValIleGlyLysMetMetAlaGlyValIle 140
404 TTCTCTCGGCACAAAGCTCTCGTCGAGAGA...TTTCTCGTGTATA 450
141 GlyGln..Phe.....LeuAlaAspProThrAspLeuValLysValGln 154
451 GCTCAGAAATGCAGATGATGCTACACCTGATTTGTCACAAAGTGA 500
155 MetGlnMetGluLysArgLysLeuGluGlyLysProLeuArgPheAr 171
501 ATCCAAACCATGTATGATGGTGAGCCGCAAGCCCTGAACCGGTCTC 550
171 gGlyValHisAlaPheAlaLysIleLeuAlaGluGlyIleArg 188
551 GGGACCATGAGAGCTTTTACCAAAATCTTACATCAGAAAGGTATAAG 600
188 ILeuTPrpAlaGlyTyrValProAsnIleGlnArgAlaAlaLeuValAsn 204
601 GGTATGGAAGAGTGTCTTCCAACATCCAGAGACATTTCTAGTAAT 650
205 MetGlyAspLeuThrTyrAspThrValLysHisTyrLeuValLeuAs 221
651 ATGGGAAACATGCTGTATGATCAAGCCAAACACTTGTCTCATGATA 700
221 nThrProLeuGluAspAsnIleMetThrHisGlyLysSerSerLeuLys 238
701 GAAGATTGCTGAGATACATTTTGGCGACACTTGTCTTATATATGT 750
238 erGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSer 254
751 CGGGTCTTCTCGACAAAGTTGATGTCACGCTGATGTGTGAAGACG 800
255 ArgIleMetAsnGlnProArgAspLysGlnGlyArgGlyLeuLeuTyr 271
801 AGATGATGAAC.....CAGGTGAATAATCTGTGTACAG 835
271 sSerSerThrAspCysLeuIleGlnAlaValGlnGlyLysPheMetS 288
836 AATTTCTTACAGACTTTTGTGAAGACGTTAAGTTGAAGGATAAGAG 865
288 erLeuTyrLysGlyPheLeuProSerTPrpLeuArgMetThrProTPrpSer 304
886 CTTTGGAAAGCTTCTTCCGACATGAGGCAAGGCTTGAACCGTGACG 935
305 MetValPheThrPheLeuThrTyrGluLysIleArgGluMetSerGlyAla 321
936 TTGCTGTGGGTCTCTATGAGAGTTAGACTGTGCGAGGAATCTC 965
321 rProPhe 323
986 TTCTCTC 992

```

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:C40226

seq_documentation_block:
ID C40226 standard; DNA: 1242 BP.
XX

```

AC C40226;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 27529.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134216.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

```

```

2  SeValPProGluGluGluGluArgLeuProLeuThrGlnArgTyrPr 18
      |||      |||      |||      |||      |||
57  TCAATCCCGCTAAATGGAGCGG...AGCCAGAGCCACAGAGAAGCACC 103
      |||      |||      |||      |||      |||
18  oArgAlaSerLysPheLeuLeuSerGlyCysAlaAlaThrValAlaGlu 35
      |      :      :      :      :      :      :      :
104  GACCGGACACAGAGATCTGCTTGGCTTCACTCTCTCCGATGGTTCAAGT 153
      |||      |||      |||      |||      |||      |||
35  euAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetGluGly 51
      :      :      :      :      :      :      :

```

154 CAGTTACGTTCCGATGTGATCTCACAAGACATGAGATGCGATGCT 203
52 GUAUAAALALEUAA.....ArgLeuGlyAspGlyAlaArgGluSe 65
204 TCGGGATCCGCTTCTGTGGCCGATCGGATGGC..... 236
65 TAlAPrOTyRArgGlyMetValArgThrAlaLeuGlyIleIleGluGlu 81
237GCTTTCGAGTCGTATCGAGA 258
82GluGlyPheLeuLysLeuTrpGluGlyValThrProAla 94
259 TTGCGCAAGAAAGAGAGAGTGTGCTCTACAAAGGCTATCTCCGGCG 308
95 ILeYrArGHisValValTyRserGlyArgMetValThrTyRGlunH 111
309 ATTATCAGACACCTGTCTACAGCGCTATCAGATCATTTGGATTGAGA 358
111 sLeuArGluValValPheGlyLysSerGlu.....AspGluHst 125
359 TTGGAAGAGACTCATGCTC...AGATCTGAAACTAACAATACGAGTCTC 405
125 yRProLeu...TrpLysSerValIleGlyGlyMetMetAlaGlyValIle 140
406 TTCCCTCGCCACAAAGGCTCTCGTGAGGA...TTTCTGTGCTTATA 452
141 GlyGln..Phe.....LeuAlaAsnProThrAspLeuValLysValGln 154
453 GGTCAAGAAATGCAAGTAGTGCTAGTCCAGCTGATTTGGTCAAACTGAGA 502
155 MetGlnMetGluGlyLysArgLysLeuGlyLysProLeuArgPheAr 171
503 ATGCAGAGAGATGATGATGTTGGTGAGCCAAAGCTGAAACCAGGTACTC 552
171 gGlyValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgG 188
553 GGGACCAATCGAGGCTTTTACCAAAATCTTACAAATCAGAGAGATGAAAG 602
188 ILeuTrpAlaGlyTrpValProAsnIleGlnArgAlaAlaLeuValAsn 204
603 GGTATGGAAGGTGTCTTCCAAACATCCAGAGAGATTTCTATGAT 652
205 MetGlyAspLeuThrThrTyRAspThrValLysHisTyRLeuValLeuAs 221
653 ATGGAGAACTAGCTGTGATGATCAGCGCAAAACATTGTATCATGARA 702
221 nThrProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuGly 238
703 GAAGATTGCTGAGATATACATTTTGGCAGACACTCTGCTTATATAGT 752
238 eRGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSer 254
753 CCGGCTTGTGCTGACAAAGTTGAGTTCGCCAGTGAATGTGGAAGAG 802
255 ArgIleMetAsnGlnProArgAspLysGlnGlyArgGlyLeuLeuTyRly 271
803 AGGATGATGAAC.....CAGGCTGAAATGCTGTATACAG 837
271 sSerSerThrAspCysLeuIleGlnAlaValGlnGlyLysGlyPheMetS 288
838 AAATCTTACGACTTTTGGTCAACACGCTTACGTTGAGAGAAATACAG 887
288 eRLeuTyRlyGlyPheLeuProSerTrpLeuArgMetThrProTrpSer 304
888 CTTTGTGGAAGAGTTTCTCCGACATGGGCAAGCTTGAGCAGGCGAG 937
305 MetValPheTrpLeuThrTyRGlulysIleArgGluMetSerGlyValSe 321
938 TTCGCTTTTGGTCTCTATGAGAAATTATGATGCTGCGAGAAATCTC 987
321 rProPhe 323
988 TTCCCTTC 994

seq_name: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.A37719
ID documentation block:
ID A37719 standard; cDNA; 876 BP.
XX
AC A37719;
XX
DT 22-NOV-2000 (first entry)
XX
DE Novel human protein coding sequence #1.
XX
KW Novel human protein; NHP; uncoupling protein; diagnosis; therapy;
KW metabolism; thermogenesis; obesity; cachexia; ss.
XX
OS Homo sapiens.
FH
FT Key Location/Qualifiers
FT CDS 1..876
FT /*lag- a
FT /*product= NHP#1
XX
XX WO200047617-A1.
XX
XX 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US03425.
XX
XX 09-FEB-1999; 99US-0119228.
XX 08-OCT-1999; 99US-0158458.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Turner CA, Mathur B, Zambrowicz B, Sands AT;
XX
XX WPI: 2000-558191/51.
XX
XX P-PSDB: Y90322.
XX
XX Nucleic acids encoding human uncoupling proteins useful for the study
XX and modulation of conditions such as obesity and cachexia -
XX
XX Claim 1; Page 25; 31pp: English.
XX
XX This sequence encodes a novel human protein (NHP) of the invention. The
XX NHP's of the invention are uncoupling proteins. The DNA and the encoded
XX protein may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate expression of the protein. The DNA
XX may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of NHP by
XX expressing inactive proteins or to supplement the patient's own production
XX of NHP polypeptides. Additionally, the DNA may be used to produce the
XX protein. Conversely, complementary sequences and antisense nucleic acid
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own NHP genes and preventing their expression. The
XX DNA and the antisense molecules may also be used as DNA probes in
XX diagnostic assays to detect and quantitate the presence of similar
XX nucleic acid sequences in samples, and hence which patients may be in
XX need of restorative therapy. They may also be used to study the
XX expression and function of the NHP's and their role in metabolism by the
XX creation of transgenic and/or gene knock-out animals. The NHP
XX polypeptides may be used as antigens in the production of antibodies
XX against NHP and in assays to identify modulators (agonists and
XX antagonists) of NHP expression and activity. Anti-NHP antibodies and
XX NHP antagonists may also be used to down regulate NHP expression and
XX activity. Anti-NHP antibodies may also be used as diagnostic agents for
XX detecting the presence of NHP polypeptides in samples. Processes and
XX conditions associated with NHP expression and activity include
XX thermogenesis, obesity and cachexia.
XX
XX Sequence 876 BP; 237 A; 184 C; 229 G; 226 T; 0 other;
XX
XX alignment_scores:

Quality: 560.00 Length: 295
Ratio: 2.605 Gaps: 4
Percent Similarity: 72.881 Percent Identity: 38.983

Alignment_block:
US-09-397-342-1 x A37719 ..

Align seg 1/1 to: A37719 from: 1 to: 876

```

23 PheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrPheArg 39
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 TTTGTGACGGGGGCTGCTCCATCATCTGCTGAGTGGGACATTTCC 74
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 OleuAspLeuThrIleuThrArgLeuGluMetGlnGluAlaAlaLeuA 56
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 AATGATTTAAACAGACACGCGCTCCAGATTCAGGCCAGC..... 117
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 IaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMetVal 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
118 .....AATGATGCAGAAATTAGGAAATTAGATACCGAGATGTG 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 ArgThrAlaLeuGlyIleGluGluGlyPheLeuLysLeuTrpGlu 69
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 CACGATAGTAGAGATAGCAGAGAGAGAGGCTGAAGACACTCTACTC 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
89 nglValThrProAlaIleTyrArgHisValValTyrSerGlyArgm 106
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210 GGGGATTCGCCCGGATTTACGCCAGGACATCCATCCATCAAGA 259
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 etValThrTyrGluHisLeuArgGluValAlaPheGlyLysSerGluAsp 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 TAGGCACTTACAGAGCTTGAAGCAGCATTTGATGAAGCCAGACAGAT 309
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 GluHisTyrProLeuTyrLysSerValIleGlyIleMetMetAlaGlyVa 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 GAAACTCTACCGATA.....AATGTGATGTGGAATTTCTGTCTGAGAT 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
139 ILeGlyIlePheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
354 CATATCTTCAACCATTTGCTTAATCCAACTGATGTTGAAAATTCGAGATC 403
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
156 ImeGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArgGly 172
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 AAGCCCA...AGCAACACATCTCAGCA.....GGA 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 ValHisHisAlaPheAlaIleLeuAlaGluGlyIleArgGlyLe 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 ATGATAGGCACTTCATGACATTTACAGCAAGAGGAGACAGAGAGACT 482
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 uTrpAlaIleTyrPvalProAsnIleGlnArgAlaAlaLeuValAsnMetG 206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
483 GTGGAAGGCTGTGCTTACTGCGCAGAGGCTGCTTATTTGTTGGTG 532
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
206 LysAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsnThr 222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
533 TGGACCTCCGCTATATGACATACCAAGAACATCTTATTTCTTCAGGC 582
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 ProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuGlySerG 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
583 CTGATGGAGACACTGTGATACCACTTCCTCAGCTTCACCTGAGG 632
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 yLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerArgI 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
633 TCTGTACAGGGGCGCTGCGCTCAACCCCTGTGATGTTGTGAGACACGTA 682
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 LemetAsnGlnProArgAspLysGlnGlyArgGlyLeuLeuTyrLysSer 272
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
683 TGATTAATCAGAGAGTGTGCGAGATGCGAGATGTTCTGCGTACACAGGA 732
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 SerThrAspCysLeuIleGlnAlaValGlnGlyGluGlyPheMetSerLe 289
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
733 ACCCTGATTTGCTGTACAGACATGAGAAAGATGAAGGTTTGTGCTCT 782

```

289 uTyrIleGlyPheLeuProSerThrLeuArgMetThrProTrpSerMetV 306
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 CTATTAAGGCTTTTGGCCAAATGTTGTTGAGACTTGTGCTTGGATATCA 832
306 alphaPheThrLeuThrTyrGluLysIleArgGluMet 317
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
833 TTTTCTTGTGACATACGACGAGCTTGAGAAATG 867

seq_name: /SID56/gcgdata/geneseq/geneseq/NA2000.DAT:289945

seq_documentation_block:
ID 289945 standard; cDNA; 936 BP.

AC 289945;

DT 05-MAY-2000 (first entry)

XX Human uncoupling protein 4 (UCP-4) nucleotide sequence.

XX Uncoupling protein 4; UCP-4; human; mitochondrial proton transport;
KW increase thermogenesis; obesity treatment; fat decrease; diabetes; fever;
KW respiratory ATP synthesis; malignant hyperthermia; ss.

XX Homo sapiens.

XX

XX Key Location/Qualifiers
XX CDS 37..909
XX /*tag= a
XX /product= UCP_4
XX /note= "Uncoupling protein 4"

XX NO200004037-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99MO-US15861.

XX 14-JUL-1998; 98US-0092737.

XX (AMYL-) AMYLIN PHARM INC.

XX Albrandt K, Beaumont K, Young AA;

XX WPI, 2000-161095/14.

XX P-PSDB; Y78511.

XX Novel uncoupling protein 4 (UCP-4) used to treat obesity, diabetes,
XX fever, malignant hyperthermia and fever -

XX Claim 5; Fig 1; 72pp; English.

XX This sequence represents the human uncoupling protein 4 (UCP-4)
XX nucleotide sequence. UCP-4 is expressed in various tissues including
XX brain, heart, pancreas and muscle tissue. UCP-4 may have a role similar
XX to other uncoupling proteins in mediating mitochondrial proton transport.
XX The UCP-4 protein has antidiabetic, anorectic and antihypertensive properties.
XX The UCP-4 protein can be used for increasing thermogenesis, treating
XX obesity, decreasing the amount of fat, or treating diabetes. Antisense
XX nucleic acids to UCP-4 are used for decreasing thermogenesis or
XX respiratory ATP synthesis. The UCP-4 protein is used to screen for
XX compounds which bind to or modulate the activity of UCP-4. Compounds that
XX increase activity of UCP-4 are used for treating conditions or disorders
XX that can be ameliorated by increasing the level of thermogenesis in a
XX subject (e.g. obesity or diabetes). Compounds that decrease UCP-4
XX activity are used for treating conditions or disorders that can be
XX ameliorated by decreasing the level of thermogenesis in a subject
XX (e.g. malignant hyperthermia or fever). Modulatory compounds can also be
XX used to regulate insulin secretion, and for increasing energy-maximizing
XX responses or increasing energy-minimizing responses.

XX Sequence 936 BP; 252 A; 183 C; 231 G; 268 T; 2 other;

CC This sequence represents a human uncoupling protein 5 (UCP5) encoding
CC DNA sequence. UCP5 is involved in metabolism, and it may be involved in
CC catalysing H⁺ leak, and therefore be involved in energetic inefficiency
CC in vivo. The present invention relates to human and murine UCP5
CC nucleotide and protein sequences. There are three isoforms of human UCP5
CC hUCP5L, hUCP5S, hUCP5SI, and two isoforms of murine UCP5, mUCP5L and
CC mUCP5S. The human UCP5 gene is located on chromosome 10q23-25. The
CC nucleic acids encoding UCP5 can be used as hybridization probes. In
CC chromosome and gene mapping, for the generation of antisense RNA and DNA
CC and in the preparation of recombinant UCP5 proteins. UCP5 nucleic acids
CC can be used in gene therapy for regulation of metabolic conditions.

Sequence 1009 BP; 267 A; 197 C; 249 G; 296 T; 0 other;

Quality:	557.50	Length:	296
Ratio:	2.617	Gaps:	5
Percent Similarity:	71.959	Percent Identity:	39.865

US-09-397-342-1 x A27996

Align seg 1/1 to: A27996 from: 1 to: 1009

23 PheLeuLeuSerGlyCysAlaAlaThrValaIaGIuLeuaIaThrPhepr 39
|||::: |||::: |||||||:::|||||
121 TTTGTATATGGCGCCCTTGCCCTATCGTGCCTAGTTTGGACTTCCC 170

39 OleuaspleuthrlysthrArgleuglmetgnglyuaalaleua 56
|:::|||||||:::|||||:::|
171 TGtGACCTTACC AAAACAGCAGCTTCAGGTTCAAGGCCAAAGCATTTGATG 220

```

56 laargleuglyaspgllyalargcluseralprotyrargglymetval 72
   ||||::: |||||
221 CCCGTTTC.....AAAGAGATAAATATACAGCGCATCGTTC 255

```

```

73 ArgThrAlaLeuGlyIleIleGluGluGluGlyPheLeuLysLeuTrpG1 89
      :::::  ::  |||  ::::| | | | | | | | | | | | | | | |
256 CATGCGCTGTTCCGATCTGTAAAGAGGAAGGTGTAATGGCTCTATTC 309

```

```

89  ngllyvalthrproalailetyrarghisvalvaltyrserglyalargm 106
      :::::::::::::::::::: :::::::::: ::::
306  AGGAATTGCTCCTGCGTTGCTAAGACAAGCATCATGCGACCATTAATA 355

```

106 etvalThrTyrgluHisLeuArgGluValValPheGlyLysSerGluasp 1222
:: ||:: ||::: :: ||||
356 TTGGATTTACCAAGCTTGAAGCGTTATTGCTAGAACGTTTAGAAGAT 4095

```

123 GlnHistryProleutRplysSerValIleGlyGlymetMetAlaGlyVa 135
    |||      |||      :::::||||| |||:::~::~:~::~::~::~::~
406 GAA...ACTCTTTAATTAATATGATCTGTGGGTAGTGTGAGGAGT 449

```

139 IIEGlyGIInPheLeuAlaAsnProThrasPheValIysValGlnMetC 156
|||::: :|||::: :|||::: :|||::: :|||::: :|||
450 GATATCTTCACACTATAGCCAAATCCACCAGATTCTTAAAGATTGCAATGC 499

```

156 lmetgluglylsrlyslengluglylsproleuargphearggly 177
    ||::|||::: |||::|||
500 AGGCTCAAGGAGC.....TTGTTCCAAGG 525

```

173 ...ValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgGlu 188
 :: ::::: || :::: ||:::
 526 AGCATGATTGGAGCTTTATCGATATATACCAACAAGAGGCACACAGGG 579

AC Z98032;

DT 09-MAY-2000 (first entry)

DE	Human secreted protein encoding nucleotide sequence SEQ ID NO:26
	<p> 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 51</p>

KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;

KW tumour; neurodegenerative disorder; developmental abnormality; allergy;

KW autoimmune disease; hepatic disease; renal disease; inflammation;

KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
asthma; sepsis; acne; psoriasis; cardiovascular disorder;

KW metabolic disorder: food additive: x chromosome: ss.
KW reproductive disorder; gastrointestinal disorder; respiratory disorder

OS Homo sapiens.

PN WO200004140-A1

PD 27-JAN-2000.

PF 14-JUL-1999; 99WO-US15849

PR 15-JUL-1998; 98US-0092921

PR 15-JUL-1998; 98US-0092956.

PA (HUMA-) HUMAN GENOME SCI INC
XX

PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
PI Lafleur NW, Ehnert P, Olsen HS, Brewer LA, Florence KA, Young BE

PI Mucenski M, Endress GA, Soppet DR;
XX

DK WFL; 2000-101126/14

PR 30-NOV-1998: 98US-0110286.
PR 16-APR-1999: 99US-0129583.
PR 15-JUL-1999: 99US-0143886.

(GETH) GENENTECH INC.

Adams S, Pan J;

WPI: 2000-412284/35.
P-PSDB: Y94669.

Isolated nucleic acid encodes human uncoupling protein 5 useful in
diagnostic assays and treatment of obesity, stroke, trauma, sepsis and
infection -
Claim 3: Fig 15: 90pp: English.

This sequence represents a murine uncoupling protein 5 isoform mUCP5S
encoding DNA sequence. UCP5 is involved in metabolism, and it may be
involved in catalyzing H⁺ leak, and therefore be involved in energetic
inefficiency in vivo. The present invention relates to human and murine
UCP5 nucleotide and protein sequences. There are three isoforms of human
UCP5, hUCP5L, hUCP5S, and hUCP5I, and two isoforms of murine UCP5, mUCP5L
and mUCP5S. The human UCP5 gene is located on chromosome 10q23-25. The
nucleic acids encoding UCP5 can be used as hybridization probes, in
chromosome and gene mapping, for the generation of antisense RNA and DNA
and in the preparation of recombinant UCP5 proteins. UCP5 nucleic acids
can be used in gene therapy for regulation of metabolic conditions.
Upregulating or downregulating UCP5 activity in a mammal is used for
modulating metabolic rate in the mammal, in particular upregulation of
UCP5 activity stimulates an increase in metabolic rate in an obese
mammal. Other therapeutic applications associated with modulating UCP5
activity are treating symptoms associated with stroke, trauma (e.g. burn
trauma), sepsis and infection. Detecting UCP5 activity can be used to
assist predictions concerning metabolic conditions or risk for onset of
obesity and as UCP5 may control the generation of reactive oxygen to
diagnose impaired neural activity or neural degeneration. Anti-UCP5
antibodies can be used in diagnostic assays and for the affinity
purification of UCP5 from recombinant cell culture or natural sources.

Sequence 1031 BP; 270 A; 199 C; 264 G; 298 T; 0 other;

alignment_scores:

Quality: 555.50 Length: 296
Ratio: 2.608 Gaps: 5
Percent Similarity: 71.959 Percent Identity: 39.865

alignment_block:

US-09-397-342-1 x A27999 ..

Align seg 1/1 to: A27999 from: 1 to: 1031

```
23 PheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrPhePr 39
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
157 TTTGTGATGCGCGCTGCTATTTGTCGCGAGTTCGCGACCTTCCC 206
39 OleuAspLeuThrIlyThrArgLeuGlnMetGlnGlyGluAlaAlaLeuA 56
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
207 TGTGAGACTTACTAAACACGCGTCAAGTCCAAAGCCAGAGTATCGATG 256
56 IaaArgLeuGlyAspGlyAlaArgIleuSerAlaProTyrArgGlyMetAl 72
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
257 TTCGTTTC.....AAAGAAATTAATATAGAGGAGGTGTTT 291
73 ArgThrAlaLeuGlyIleIleGluGluGluPheLeuLysLeuTyrGcl 89
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
292 CATGCGCTTGTTCGAGATCTATAAGAGAGAGGATCTTGCTGTATTC 341
89 nglyValThrProAlaIleTyrArgHisValValTyrSerGlyGlyArgm 106
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
342 AGGAATTCGCCCTGCTACTAAGACAGCATCATATGACACATCAAAA 391
```

```
106 etValThrTyrGluHisLeuArgGluValValPheGlyLysSerGluuAsp 122
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
392 TTGCTATTATGAAAGCTTGAAGCATATTGTTAGAACGTTTGGAAAGAT 441
123 GluHisTyrProLeuTyrPlySerValIleGlyLysMetAlaGlyValA 139
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
442 GAG.....ACTCTCTAATTAACATGATCTGTGGGGTAGTCAGAGAGT 485
139 IileGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGlnMetG 156
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
486 GATTTCCTCTACTATTCGCAATCCCACTGATGTTCTAAAGATTCGAATGC 535
156 ImetGluGlyLysArgLysLeuGluGlyLysProLeuArgPheArgGly 172
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
536 AGGCTCAAGGAGT.....TTGTTCCAGGG 561
173 ..ValHisHisAlaPheAlaLysIleLeuAlaGluGlyIleArgG1 188
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
562 AGCATGATTTGGCAGCTTCATGACATATACACAGAGAGGATACAGGGG 611
188 yLeuThrAlaGlyTyrPvalProAsnIleGlnArgAlaAlaLeuValAsm 205
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
612 TCTGTGGAGGGGTGTGTGCTCCCACTGCTCAGCTGCTCAATCGTTGTGG 661
205 etGlyAspLeuThrThrTyrAspThrValLysHisTyrLeuValLeuAsn 221
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
662 GAGTAGACCTCCCGCTTATGATATTTACCAAGAAACCTGATAGTTTCA 711
222 ThrProLeuGluAspAsnIleMetThrHisGlyLeuSerSerLeuCysSe 238
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
712 GGAATGCTGGGAGACACAAATTTAAACACATTTGTTCCAGTTTCACCTG 761
238 rGlyLeuValAlaSerIleLeuGlyThrProAlaAspValIleLysSerA 255
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
762 TGGTTTGGCTGGGCTCTGGCATCAACCTGTGATGTGTGAGAACTC 811
255 rGileMetAsnGlnProArgAspLysGlnGlyArgGlyLeuLeuTyrLys 271
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
812 GAATGATATATCAG..AGGCAATAGTGGACATGTGGACTCTACACAG 858
272 SerSerThrAspCysLeuIleGlnAlaValGlnGlyGluGlyPheMetSe 288
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
859 GGTACTTTGGATGATATTTTAAAGATGTGAAGCATGAGGATTTTTCG 908
288 rLeuTyrLysGlyPheLeuProSerThrLeuArgMetThrProTyrSerM 305
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
909 ACTTATATAAGGATTTTGGCCAAACTGGCTTCGACTTGGACCTGGACACA 958
305 etValPheThrLeuThrTyrGluLysIleArgGluMet 317
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
959 TCATTTTTTTATTTACTATGACGACTCAAGAGGCTT 996
```


OM of: US-09-397-342-1 to: Issued_Patents_NA:* out_format : pfs
Date: Apr 28, 2001 7:35 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09397342/unat_27042001_170000_22911/app.query.fasta.1.384
-DB=Issued_Patents_NA -QPR=fastap -SUFFIX=p2n.rni
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=bloms62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09397342.ECIGN1_1_59 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPEX
WAIT -THREADS=1

Search information block:

Query: US-09-397-342-1
Query length: 323
Database: Issued_Patents_NA:*
Database sequences: 302621
Database length: 87301344
Search time (sec): 65.240000

score list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-1 +			475.50	960.08	5.3e-47 1192
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-518-878B-38 +			472.50	973.07	1.3e-46 1255
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-38 +			472.50	973.07	1.3e-46 1255
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-38 +			472.50	973.07	1.3e-46 1255
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-466-1 +			472.50	970.17	1.9e-46 1596
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-210-681A-38 +			472.50	970.17	1.9e-46 1596
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-38 +			472.50	970.17	1.9e-46 1596
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-466-3 +			472.50	967.76	2.6e-46 1949
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-172-528-3 +			472.50	967.76	2.6e-46 1949
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-318-199-3 +			472.50	957.08	1.0e-45 2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-466-1 +			469.50	957.08	1.0e-45 2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-318-199-1 +			469.50	957.08	1.0e-45 2782
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-36 +			396.50	811.74	1.3e-37 1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-37 +			396.50	811.74	1.3e-37 1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-36 +			396.50	811.74	1.3e-37 1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-36 +			396.50	811.74	1.3e-37 1205
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-937-466-5 +			376.50	764.46	5.4e-35 1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-172-528-5 +			376.50	764.46	5.4e-35 1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-318-199-5 +			376.50	764.46	5.4e-35 1777
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-188-930-262 +			231.50	460.69	4.5e-18 1177
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-961-871-9 +			230.50	457.75	6.3e-18 1299
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-961-871-11 +			227.50	448.51	2.1e-17 1594
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-234-613-61 +			227.50	448.51	2.1e-17 1594
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-933-750C-68 +			220.50	433.24	1.5e-16 1643
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-234-613-68 +			220.50	433.24	1.5e-16 1643
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-5 +			173.00	340.09	2.3e-11 847
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-3 +			173.00	249.79	2.5e-06 222
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-878B-14 +			123.00	245.80	4.2e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-14 +			123.00	245.80	4.2e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-14 +			123.00	245.80	4.2e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-14 +			123.00	245.80	4.2e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-210-681-14 +			123.00	245.80	4.2e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-14 +			123.00	245.80	4.2e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-15 +			122.50	244.73	4.8e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-294-522B-15 +			122.50	244.73	4.8e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-807-861A-15 +			122.50	244.73	4.8e-06 309

/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-470-868A-15 + 122.50 244.73 4.8e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-210-681-15 + 122.50 244.73 4.8e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-946-719A-15 + 122.50 244.73 4.8e-06 309
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-188-930-23 + 108.00 199.72 0.0015 997
/cgn2_6/ptodata/1/lna/5A.COMB.seq:US-08-742-273-1 + 93.50 161.84 0.1983 1782

seq_name: /cgn2_6/ptodata/1/lna/5A.COMB.seq:US-09-142-565-1

seq_documentation block:

Sequence 1, Application US/09142565A
Patent No. 6187560
GENERAL INFORMATION:
APPLICANT: Lee James Bealey
APPLICANT: Kelly Paine
APPLICANT: Robert James
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GH-30002
CURRENT APPLICATION NUMBER: US/09/142,565A
CURRENT FILING DATE: 1999-06-30
EARLIER APPLICATION NUMBER: 9704551.2
EARLIER FILING DATE: 1997-03-05
EARLIER APPLICATION NUMBER: 9705614.7
EARLIER FILING DATE: 1997-03-18
EARLIER APPLICATION NUMBER: 97305305.1
EARLIER FILING DATE: 1997-07-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1192
TYPE: DNA
ORGANISM: HOMO SAPIEN
US-09-142-565-1

alignment_scores:

Quality: 475.50 Length: 316
Ratio: 2.331 Gaps: 7
Percent Similarity: 64.557 Percent Identity: 35.759

alignment_block:

US-09-397-342-1 x US-09-142-565-1 ..

Align seg 1/1 to: US-09-142-565-1 from: 1 to: 1192

20 AAlserlyspheuleuleSerGlycAlaAlaThraValAlaGluLeuAl 36
GCTGTGAGTTCCTGGGGGCGAGCACACGCTGTTTGTGCTGACCTGT 287
36 aThpPheleuaspheuleThraValAlaGluLeuAlaGluLeuAl 52
TACCTTCCACTGAGCACACGAGGTCCTGCTGACCTGAGGAGCA 337
53AlaAlaLeuAlaArgLeuGlyAspGlyAlaArgGluSerAla 66
ACGAGGGGCTCCACAGCGCGGCTC.....GTG 366
338 ACAGGGGCTCCACAGCGCGGCTC.....GTG 366
67 ProTyrArgGlyMetValArgThraAlaGluLeuAlaGluLeuAl 83
TACCTTCCACTGAGCACACGAGGTCCTGCTGACCTGAGGAGG 416
367 CAGTACCGTGGCTGGCGGCGAGCACATCTGACCACTGGTGGAGG 416
83 yPheLeuLysLeuThrArgGlyValAlaThrProAlaIleTyrArgHisVal 100
TTCCTGCAAGCCCTTACATGGCTGGGCGGCGGCTGAGCGGAGG 466
417 TCCCTGCAAGCCCTTACATGGCTGGGCGGCGGCTGAGCGGAGG 466
100 aTyrSerGlyArgMetValThrTyrGluHisLeuArgGluValAl 116
GCTTGGCTGCTGATCGATCGATCGATCGATCGATCGATCGATCGAT 516
467 GCTTGGCTGCTGATCGATCGATCGATCGATCGATCGATCGATCGAT 516
117 PheLeuLysSerGlyAspGluHisTyrProLeuTyrLysSerValIleG 133
ACCCCAAGGCGGCGGAGC...AGCTCCAGCTCTACCTGAGGATTTGGC 563
517 ACCCCAAAGGCGGCGGAGC...AGCTCCAGCTCTACCTGAGGATTTGGC 563
133 yGlyMetLeuAlaGlyValIleGlyGlnPheLeuAlaAsnProThrAsp 150

```

564 CGCGTCGACACACAGGAGCCATGGCGGTACCTGTGGCCAGCCACAGATG 613
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
150 euVallysValGImetGImetGlu.....GlyLysArgLysLeu 163
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
614 TGGTAAAGTCCGAATTCAGGCGCAGCTACACCTCGCGGCATCCAGAGAC 663
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
164 GlnGlyLysProLeuArgPheArgGlyValAlaHisLeuPheAlaLys 180
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
664 GACAGAAAA.....TACAGCGGAGCTATGAGACCGCTACGAAACAT 704
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
180 eLeuAlaGlnGlyGlyLeuArgGlyLeuThrPalaglyTrpValProAsn 197
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 CGCCAGGAGGAAGAGCTCAGCGGCTGTGGAAGAGAACTTGGCCACAA 754
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
197 legAlaArgAlaAlaLeuValAlaSerMetGlyAspLeuThrThrTrpAsp 213
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 TCATAGAGATCTCATCGTCAACTGTGCTGAGAGTGGTAGCTACGACATC 804
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
214 ValLysHisTrpLeuValLeuAsnThrProLeuGluAspAsnIleMet 230
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
805 CTCAGAGGAGAACTGCTGTGACTACCACTGCTCAGTACGACAACTCCCG 854
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGly 247
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
855 CCACCTTGTCTGCTGCTTGGAGCGGCTTGTGTGCTGCACAGAGTGCGCT 904
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
247 hrProAlaAspValIleLysSerArgIleMetAsnGlnProArgAsp 263
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
905 CCCCCTGAGACTGGTGAAGACCCGGTATATGATACCTACCTCCAGCCAG 954
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
264 GlnGlyArgGlyLeuLeuThrLysSerSerThrAspCysLeuIleGln 280
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
955 .....TACTTCAGCCCCCTCGACTGTATGATAAAGAT 986
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
280 aValGlnGlyGlnGlyPheMetSerLeuThrLysGlyPheLeuProSer 297
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
987 GTGGCCCCCAGAGGGCCCCACAGCCCTTCTACAAAGGATTTACACCTCCT 1036
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
297 rPLeuArgMetThrProTrpSerMetValPheTrpLeuThrTyrgLys 313
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1037 TTTTCGTTTGGGATCTCGAAGCGTGGTATGCTGTACTATGAGCAG 1086
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
314 IleArg.....GluMetSerGlyValSerProPhe 323
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1087 CTGAAGCGGGCCCTGATGAAGTCCAGATGTACGGGAATCACCGTTT 1134
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
--name: /cgn2_6/plodata/1/lna/5A_COMB.seq:US-08-518-878B-38
1. documentation block:
2. Sequence 38, Application US/08518878B
3. Patent No. 5702902
4. GENERAL INFORMATION:
5. APPLICANT: Tartaglia, Louis A.
6. TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
7. TITILE OF INVENTION: TREATMENT OF BODY WEIGHT DISORDERS,
8. NUMBER OF SEQUENCES: 57
9. CORRESPONDENCE ADDRESS:
10. ADDRESSEE: Pennie & Edmonds
11. STREET: 1155 Avenue of the Americas
12. CITY: New York
13. STATE: New York
14. COUNTRY: U.S.A.
15. ZIP: 10036-2711
16. COMPUTER READABLE FORM:
17. MEDIUM TYPE: Floppy disk
18. COMPUTER: IBM PC compatible
19. OPERATING SYSTEM: PC-DOS/MS-DOS
20. SOFTWARE: PatentIn Release #1.0, Version #1.30
21. CURRENT APPLICATION DATA:
22. APPLICATION NUMBER: US/08/518,878B
23. FILING DATE: 23-AUG-1995
24. CLASSIFICATION: 435

```

[illegible]

518 TTGCTCGTAATGCCATTGTCAGCTGCTGAGCTGTGACCTATGACCTC 567
 214 ValylshistyleuValleuAsnthrProleuLuspaanilemeth 230
 568 ATCAAGATGCTCTGTAAGACCACTGACAGATGACCTGCTCCTTG 617
 230 rHisGlyLeuSerSerLeuYsSerGlyLeuValAlaSerlleuGlyT 247
 618 CCACTTACTTCTGCTTGTGGGAGGCTTGTGACCACTGCTATGCTCCT 667
 247 hrProAlaSpValIleLysSerArglleMetasnglnProArgAspLys 263
 668 CCCCTGTAGAGCTGTGTAAGACGAGATACATGACTGCTGCCCTGGCCAG 717
 264 GlnglyArgGlyLeuLeuTyrlYsSerSerThrAspGlyLeuIleGlnAl 280
 718TACAGTAGCGCTGGCCACGCTGCTGCTTACCCT 749
 280 aValGlnGlyGlyLeuPheMetSerLeuTyrlYsGlyPheLeuProSerT 297
 750 GCTCCAGAGAGAGAGGCGCCGAGCTTCTACAAAGGTTGATGCTCCTCT 799
 297 rPleuArgMetThrProTrpSerMetValPheTrpLeuThrTyrlGlyLys 313
 800 TTCCTCCCTTGGGTTCTGGAACGTGTGATGTTGTCACCTATGAGCAG 849
 314 IleArg 315
 850 CTGAAA 855

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-294-522B-38

seq_documentation_block:
 ; Sequence 38, Application US/08294522B
 ; Patent No. 5741666
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; TITLE OF INVENTION: Compositions and Methods for the
 ; TITL OF INVENTION: Treatment of Body Weight Disorders, Including Obesity
 ; NUMBER OF SEQUENCES: 48
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/294,522B
 ; FILING DATE: 23-AUG-1994
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7853-015
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1255 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-294-522B-38

alignment_scores: Quality: 472.50 Length: 302

Ratio: 2.316 Gaps: 6
 Percent Similarity: 67.550 Percent Identity: 34.437
 alignment_block:
 US-09-397-342-1 x US-08-294-522B-38 ..

Align seq 1/1 to: US-08-294-522B-38 from: 1 to: 1255

18 ProArgAlaSer...LysPheLeuLeuSerGlyLysAlaAlaThrValAl 33
 1 CCRCTGCTGCTGTAAGTTCTTGGGGCTGGCACACCTGCTTCATCGC 50
 33 aGluLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuIleMetG 50
 51 AGATCTATCATCCTTCTGCTGCTGATGCTGTAAGTCCGGTTACAGATCC 100
 50 lnglyGlnAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63
 101 AAGGAGAAAGT.....CAGGGGCCAGTGGCGGCTACA 132
 64 GluSerAlaProTyrlArgGlyMetValArgThrAlaLeuGlylleIleG 80
 133 GTCAGCGCCCGACGTACCGGGTGTGATGGGACCATCTGACCATGTTGCG 182
 80 uGlnGlyGlyPheLeuLysLeuTrpGlnGlyValThrProAlaIleTyra 97
 183 TACTGAGGGCCCCGGAAGCTCTTACAAATGGGCTGTGGCCGCTGAGC 232
 97 rGHisValValTyrlSerGlyLysArgMetValThrTyrlGlnHisLeuArg 113
 233 GCCAAATGAGCTTTGGCTTCTGCGCATCGCGCTGTATGATTCGTCAAA 282
 114 GluValValPheGlyLysSerGluAspGlnHisTyrlProLeuTrpLysSe 130
 283 CAG.....TTCCTACACCAAGGCTGTGAGCATGCCAGCATGGAGACCG 326
 130 rValIleGlyGlyMetAlaGlyValIleGlyGlnPheLeuAlaAsp 147
 327 CTTCTTGCAGGACGACACACAGGCTGCTGCTGCTGCTGCTGCTGCTG 376
 147 rThrAspLeuValLysValGlnMetGlnMetGlnGlyLysArgLysLeu 163
 377 CCAGGAGATGCTGTAAGGTCCGATTCAGAGCTCAGGCCG..... 417
 164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLysII 180
 418 GCTGAGAGTGTGCTGAGATACCAAGACCGCTCATGCTCTACAAGACCAT 467
 180 eLeuAlaGlnGlyLysIleArgGlyLeuTrpAlaGlyTrpValProAsnI 197
 468 TGCCCCGAGAGAGAGGTTCCGGGCTCTTGGAAAGGACCTCTCCCATG 517
 197 leGlnArgAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrlAspThr 213
 518 TTGCTCGTAATGCCATTGTCAACTGTGCTGAGCTGTGAGACTATGACTC 567
 214 ValylshistyleuValleuAsnthrProleuLuspaanilemeth 230
 568 ATCAAGATGCTCTGTAAGACCACTGACAGATGACCTGCTCCTTG 617
 230 rHisGlyLeuSerSerLeuYsSerGlyLeuValAlaSerlleuGlyT 247
 618 CCACTTACTTCTGCTTGTGGGAGGCTTGTGACCACTGCTATGCTCCT 667
 247 hrProAlaSpValIleLysSerArglleMetasnglnProArgAspLys 263
 668 CCCCTGTAGAGCTGTGTAAGACGAGATACATGACTGCTGCCCTGGCCAG 717
 264 GlnglyArgGlyLeuLeuTyrlYsSerSerThrAspGlyLeuIleGlnAl 280
 718TACAGTAGCGCTGGCCACGCTGCTGCTTACCCT 749
 280 aValGlnGlyGlyLeuPheMetSerLeuTyrlYsGlyPheLeuProSerT 297

```

750 GCTCCAGGAGGGGCCCGACCTTTCTTACAAGAAGGTTCAGCCCTCCT 799
      ||||| ||||| :|||||
297 rpleuarqmetrhrrProTrPserMeValPhehrleuthrlyrgulys 313
      :::|||||::: ||::::|||::::|||::::|||::::
800 YTCCTCCGTGGTTCCTGTGAACGTGTATGTCCTCATATGAGCAG 849
      314 lIearg 315
           ::|||
850 CTGAAA 855

seq_name: /cgn2_6/ptodata/1/ina/5B-COMB.seq:US-08-470-868A-38

seq_documentation_block:
; Sequence 38, Application US/08470868A
; Patent No. 5861485
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis C.
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Treatment of Body Weight Disorders, Including Obesity
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,868A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-0031-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-470-868A-38

alignment_scores:
Quality: 472.50 Length: 302
Ratio: 2.316 Gaps: 6
Percent Similarity: 67.550 Percent Identity: 34.437

Alignment_block:
US-09-397-342-1 x US-08-470-868A-38 ..

Align seg 1/1 to: US-08-470-868A-38 from: 1 to: 1255

18 ProAghAlaser..LysPheLeuleuSerGlyCySAIAalThVal 33
   || |::|::| ||||| ||::|| ||||| ::||
1 CCTACTGCACCTGTGAAGTTCTTGCGCGTGCAACAGCTGCCTGCATCC 50
3 aglueulaatArpheProLeuAspleuThyLysThrArgleuInMet 50
   ::||| ||||| ||||| ||::||| ||||| |||||
51 AGATCTCAACACCTTCCCTCTGTGATCTCTAAAGTCCGATTACAGATCC 100
50 InglyeLuAlaAlaLeuAlaArgleugLyAspGly.....AlaArg 63

```

```

101 AAGGAGAACT.....CAGGGCCACTGGCGCTTACA 132
64 GUSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleIlecl 80
133 GTGAGGCCCCAGTACCGGGTGTGTGGGCACCAATTCTGCACCATGGTGGC 182
80 uGluGluGlyPheLeuLysLeuTrpGluGlyValThrProAlaIleTyrA 97
183 TACTGAGGGCCCCGAAACCTCTACAAATGGGCTGGTGGCGGCTTCACG 232
97 rGHisValValTyrSerGlyArgMetValThrTyrGluHisLeuArg 113
233 GCCAAATAGACTTGGCTCTCTCCGACATCGGCGCTATGATTCCTGCATA 282
114 GluValValPheGlyLysSerGluAspLinhTyrProLeuTrpLysSe 130
283 CAG.....TTCTACACCAAGGCGCTCTAGACATGCCACATTGGGAGCCG 326
130 ValIleGlyIleMetMetAlaGlyValIleGlyLinhPheLeuAlaAsp 147
327 CCTGCTTAGCAGGACGACACCAAGGAGCCCTGGCTGGCTGGTGGCCAC 376
147 rGHisAspLeuValLysValGlnMetGlnMetGlyLysArgLysLeu 163
377 CCACGGATGGTAAAGTCCGATTCACAAAGCTCAGGCGCG..... 417
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLysI 180
418 GCTGGAGGCTGTGGAGATACCAAGACCGCATGCTCTACAAAGACAT 467
180 eleuAlaGluGlyIleArgGlyLeuTrpAlaGlyTyrValProAsnI 197
468 TGCCCGAGAGAAAGGTTCCCGGGCTCTGGAAAGGAGCCTCTCCCAATG 517
197 legIArgAlaAlaLeuValAsnMetGlyAspLeuThrThyTrpAspThr 213
518 TTGCTCTAAATGCCATTGTCAACTGTGCTGAGCTGTGACTGATGACCTC 567
214 ValLysHisTyrLeuValIleuAsnThrProLeuGluAspAsnIleMet 230
568 ATCAAGAGATGGCCCTCGAAGAACCACTCATGACAGATGACCTCCCTTG 617
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGlyT 247
618 CCACTTCACCTTGCCTTGGGGGAGGCTTGCACCACTGTCATCGCCT 667
247 hrProAlaAspValIleLysSerArgIleMetAsnGluProArgAspLys 263
668 CCCCTGAGACGTGTCAAGACGAGATCATGAACTGTCCCTGGGCGCAG 717
264 GlnGlyArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAl 280
718 .....TACAGTAGCCCTGGCCACTGTGCCCTACCAT 749
280 aValGlnGlyGluGlyPheMetSerLeuTyrLysGlyPheLeuProSerT 297
750 GCTTCCAGAAAGAGGGGCCGAGCGCTTCTACAAAGGTTCAATGCCCTCT 799
297 rPheuArgMetThrProTrpSerMetValPheTrpLeuThrTyrGlyLys 313
800 TTCTCCGCTTGGGTCTCTGGAACGTGGTGAGTTCCTGCACCATATGACG 849
314 IleArg 315
850 CTGAAA 855
seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:us-08-807-861A-38
seq_documentation_block:
; Sequence 38, Application US/08807861A
; Patent No. 5853975
; GENERAL INFORMATION:

```

APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/807,861A
FILING DATE: 26-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/518,878
FILING DATE: 23-AUG-1995
APPLICATION NUMBER: US 08/470,868
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/294,522
FILING DATE: 23-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-807-861A-38

alignment_scores:
Quality: 472.50 Length: 302
Ratio: 2.316 Gaps: 6
Percent Similarity: 67.550 Percent Identity: 34.437

alignment_block:
US-09-397-342-1 x US-08-807-861A-38 ..

Align seg 1/1 to: US-08-807-861A-38 from: 1 to: 1596

```
18 ProArgAlaSer..LysPheLeuLeuSerGlyCysAlaAlaThrValAl 33
   ||| |||..... ||| ||| |||..... ||| ||| |||..... |||
343 CCTACTGCACTGTGAAGTTCTTGGGGGTGGACACGCTGCGTCGATCGC 392
33 agtLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetG 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 AGATCTCATCACTTCTCTGTGATACGCTAAAGTCCGCTACAGATCC 442
50 lnglGluAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63
   ||| ||| |||..... ||| ||| |||..... ||| ||| |||.....
443 AAGGCAAGAGT.....CAGGGCCAGTGGCGCGCTACA 474
64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyIleLeu 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 GTCACGGCCCACTACCGGGGTGTGATGGGACCATTCGATGACCATGGT 524
80 uGluGluGlyPheLeuLysLeuTrpGlnGlyValThrProAlaIleTyr 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
525 TACTAGAGGCCCCGAGAGCCTCTACATGGGCTGGTGGCGGCTGCAGC 574
97 rghIsvaValTyrSerGlyArgMetValThrTyGlnHisLeuArg 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
575 GCCAATGAGCTTCCCTGTGCCACATCGGCGCTGATGATCTGTCAAA 624
114 GluValValPheGlyLysSerGlnAspGlnHisTyrProLeuTrpLys 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
625 CAG.....TTCTACACCAAGGCTGTGAGCATGCGACCATGGAGCCG 668
130 rValIleGlyGlyMetValGlyValIleGlyGlnPheLeuAlaAsp 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 CCTCTAGAGGACACACACAGATGCGCTGCTGCTGGTGGCGGCGCAG 718
147 rGthAspLeuValLysValGlnMetGlnMetGlyLysArgLysLeu 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
719 CCACGATGTGTAAAGTCGATTCACAGGCTCAGCGCCG..... 759
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
760 GCTGGAGGTGTGGAGATCCAAAGACCGTCAATGCTTACAGACCAT 809
180 eLeuAlaGluGlyGlyLeuArgGlyLeuThrPalaGlyTrpValProAsn 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
810 TCCCGGAGAGGAAGGTTCGGGGCTGTGAAAGGACCTTCCCAATG 859
197 lGclArgAlaAlaLeuValAsnMetGlyAspLeuThrTyTrpAspThr 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 TTGCTGTATATGCAATGTCACTGTGAGCTGTGATGATGACCTC 909
214 ValLysHisTyrLeuValLeuAsnThrProLeuGlnAspAsnIleMet 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
910 ATCAAGATGCTCTCTCAACCACTCATGACATGATGACCTCTTG 959
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGly 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
960 CCACCTCACTTGTGCTTGGGGAGGCTTGTGACACCATGTGATGCCCT 1009
247 hProAlaAspValIleLysSerArgIleMetAsnGlnProArgAspLys 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1010 CCCCTAGACGTGTCAAGACAGATATCATGATCTGCTGCGGCGCAG 1059
264 GlnGlyArgGlyLeuLeuTyrLysSerThrAspCysLeuIleGlnAl 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 .....TACAGTACCGCTGCCACTGTGCTTACCAT 1091
280 aValGlnGlyGlyGlyPheMetSerLeuTyrLysGlyPheLeuProSer 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 GCTCCAGAGAGAGGGCCCCGAGCCTTCTACAAAGGTTCAATGCCCTCT 1141
297 rPLeuArgMetThrProTrpSerMetValPheTrpLeuThrTyGlnLys 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1142 TTCCTCGCTTGGGTCTCTGGAACGTGTGATGTCTGCTACCTATGAGCAG 1191
314 lLeuArg 315
1192 CTGAAA 1197
seq_name: /cgn2_6/plodata/1/ina/6a_COMB.seq:us-09-210-681-38
seq_documentation_block:
Sequence 38. Application US/09210681
Patent No. 6057109
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
```

```

? COUNTRY: U.S.A
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent in Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/210,681
? FILING DATE:
? CLASSIFICATION:
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 08/807,861
? FILING DATE: 26-FEB-1997
? APPLICATION NUMBER: US 08/518,878
? FILING DATE: 23-AUG-1995
? APPLICATION NUMBER: US 08/470,868
? FILING DATE: 06-JUN-1995
? APPLICATION NUMBER: US 08/294,522
? FILING DATE: 23-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Coruzzi, Laura A.
? REGISTRATION NUMBER: 30,742
? REFERENCE/DOCKET NUMBER: 7853-066
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-9741/8864
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 38:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1596 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? US-09-210-681-38

alignment_scores:
? Quality: 472.50 Length: 302
? Ratio: 2.316 Gaps: 6
? Percent Similarity: 67.550 Percent Identity: 34.437

alignment_block:
US-09-397-342-1 x US-09-210-681-38 ..
Align seg 1/1 to: US-09-210-681-38 from: 1 to: 1596

18 ProArgLaser...LysPheLeuLeuSerGlyCysAlaAlaThrValAl 33
   ||| |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
343 CCTACTGCGACCTGTGAAGTTCTTGGGGCTGGCACACCTGCGTCATCGC 392
33 agLeuAlaThrPheProLeuAspLeuThrLysThrArgLeuGlnMetG 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 AGATCTCATACCTTCTCTGCTGACTGCTAAAGTCCGGTACAGATCC 442
50 lngLysAlaAlaLeuAlaArgLeuGlyAspGly.....AlaArg 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 AAGGAGAAAGT.....CAGGGGCGACGTGGCGGTACAA 474
64 GluSerAlaProTyrArgGlyMetValArgThrAlaLeuGlyLeuIleG 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 GTGAGCGCCGAGTACCGGGGTGATGGGACCATTCGTGACATGGTGGC 524
80 uGluGluGlyPheLeuLysLeuTyrGlnGlyValThrProAlaIleTyr 97
   : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
525 TACTGAGGCGCCCGAAGCTTACATAGGCTGGTGGCGGCTTGACG 574
97 rGHisValValTyrSerGlyArgMetValThrTyrGlnHisLeuArg 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
575 GCCAAATGACCTTGGCTGTCCGATCGCGCTGATGATGATGTCGACAA 624
114 GluValValPheGlyLysSerGluAspGlnHisTyrProLeuTyrPyl 130

```

```

625 CAG.....TTCTACACCAAGGCTGTGACATGCGACGATTCGGACCG 668
130 rValIleGlyGlyMetMetAlaGlyValIleGlyGlnPheLeuAlaAsp 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 CTCCTGACGAGGAGGACGACACAGGTGCGCTGCTGCTGCTGCTGCTG 718
147 rothAspLeuValLysValGlnMetGlnMetGluGlyLysArgLysLeu 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
719 CCAAGATGTGTAAAGTCCGATTCACAGCTCAGCGCCG..... 759
164 GluGlyLysProLeuArgPheArgGlyValHisHisAlaPheAlaLys 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
760 GCTGAGAGGTGCTGCGAGATACCAAGGACCGTCAATGCTTACAGACCA 809
180 eLeuAlaGluGlyGlyLeuArgGlyLeuTyrPalaGlyTyrPalaPro 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
810 TGCCCGAGAGGAGGTTCCGGGCTCTGTGAAAGGACCTTCCCATG 859
197 leGlnArgAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrAspThr 213
   : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 TTGCTGCTAATGCCATGTCACTGTGCTGAGCTGCTGAGCTATGACCTC 909
214 ValLysHisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMet 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
910 ATCAAGATGCTCCCTCTGAAAGCCACCTCATGACATGATCCCTTGG 959
230 rHisGlyLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGly 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
960 CCACCTTACTTCTGCTTGGGGAGGCTTCTGACACCATGCTATGCTCT 1009
247 hrProAlaAspValIleLysSerArgIleMetAsnGlnProArgAspLys 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1010 CCCCTGATGACGTGTCACAGCAGATACATGAATCTGCTGCGGCGAC 1059
264 GluGlyArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGln 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 .....TACAGTACGCTGGCCACTGCTGCTTACCAT 1091
280 aValGlnGlyGlyGlyPheMetSerLeuTyrLysGlyPheLeuProSer 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 GCTCCAGAGAGGAGGCGCCGACCTTCTACAAAGGTTATGCTCTCT 1141
297 rPLeuArgMetThrProTyrSerMetValPheThrPheThrTyrGluLys 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1142 TTCTCCGCTTGGGTCTCTGGAAGTGTGATGTTCTGACCTATGAGCAG 1191
314 IleArg 315
1192 CTGAAA 1197

seq_name: /cgn2_6/ptodata/1/lna/6B_COMB.seq:US-08-946-719A-38
seq_documentation_block:
? Sequence 38, Application US/08946719A
? Patent No. 6121017
? GENERAL INFORMATION:
? APPLICANT: Tartaglia, Louis A.
? TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
? TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
? NUMBER OF SEQUENCES: 64
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pennie & Edmonds LLP
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent in Release #1.0, Version #1.30

```

[illegible]

```

: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: T97-009
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 343-4341
: TELEFAX: (650) 343-4342
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1949 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-09-397-466-3

```

```

alignment_scores:
  Quality: 472.50      Length: 300
  Ratio: 2.386         Gaps: 7
  Percent Similarity: 66.000  Percent Identity: 35.667

```

Alignment_block:

US-09-397-342-1 x US-08-937-466-3

Align seg 1/1 to: US-08-937-466-3 from: 1 to: 1949

```

22  LysPheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrph 38
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
285  AAGTTCCTGGGGCGGCGGCGGCGGCTGTTTGGGAGCTCTCCACCTT 334
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
38  eProlLeuLeuThrThrThrThrThrThrThrThrThrThrThrThrThr 55
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
335  TCCCTGGACACCGGCAAGGTCGCTGCAGATCCAAAGGGGAG..... 377
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
55  eAlaIArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMet 71
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
378  .....AACCCAGGGGCT...CAGAGCGTGCAGTACCCGGGTGTG 413
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
72  ValArgThrAlaLeuGlyIleIleGluGluGluGluGluGluGluGlu 88
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
414  CTGGGTACCTCCTGATGATGTCGCGCAGAGGGTCCCGGAGCCCTTA 463
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
88  pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
464  CAGGGAGCTGTGCTGGCTGGACGCGCAGATGAGTTTGGCTCCATTC 513
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
105  rGmetValThrTyrGluHisLeuArgGluValValPheGlyLysSerGlu 121
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
514  GAATTGGCCCTTAAGACTGTGTCAAGCAGTTTACACCCCGGAGGAGCG 563
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
122  AspGluHisTyrProLeuTyrLysSerValIleGlyGlyMetMetAla 138
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
564  GAC...CACTCCAGCGTCGCCATCAGGATTCGTGCAAGGCTGCACAG 610
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
138  yValIleGlyGlnPheLeuAlaAsnProThrAspLeuValValGln 155
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
611  AGCCATGGCAGTGAACCTGCGCCAGCCAGAGATGTGGTGAAGGTCGAT 660
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
155  eGln.....MetGluLysArgLysLeuGluGly 165
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
661  TTCAGGCGCATGATACGCTGGAACTGGAGAGAGAGAGAA..... 701
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
166  LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeu 182
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
702  .....TACAGAGGAGCTATGAGCTTACAGAACCATCGCCAG 739
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
182  aGluGlyGlyIleArgGlyLeuTyrPheAlaGlyTyrValProAsnIle 199
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
740  GAGGAGAGAGTCAAGGCGCTGTGGAAAGGAGCTTGGCCCAACATCAAA 789
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
199  rGAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrAspThrVal 215
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
790  GAATGCGCATGTCAACTGTCTGAGATGTGTGACCTACGACATCATCAAG 839
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

216  HisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHis 232
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
840  GAGAACTTCTGCTGAGTCTCACTGTTTACAGAACTTCCCTGTGACT 889
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
232  yLeuSerSerLeuCysSerGlyLeuValAlaSerIleLeuGlyThrPro 249
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
890  TGCTCTGCTTGGAGCTGTGCTGTGTGCTGCTGCTGCTGCTGCTGCT 939
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249  LaaPValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGly 265
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
940  TGGATGTGTAAACACCGATACATGAACTCC.....CTAGGC 980
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
266  ArgGlyLeuLeuTyrLysSerThrAspCysLeuIleGlnAlaValAla 282
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
981  AGG.....TACGCGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
282  nGlyGluGlyPheMetSerLeuTyrLysGlyPheLeuProSerThrLeu 299
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1022  TCAAGAGGAGCCACAGGCGCTTCAAGAGATTTGTGCTCCCTTCTGCT 1071
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

seq_name: /cgn2_6/plodata/1/lna/5B_COMB.seq:US-09-172-528-3

seq_documentation_block:

```

: Sequence 3, Application US/09172528
: Patent No. 5952469
: GENERAL INFORMATION:
: APPLICANT: Zhang, Ning
: APPLICANT: Amaral, M. Catherine
: APPLICANT: Chen, Jin-Long
: TITLE OF INVENTION: UCP3 Genes
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 75 DENISE DRIVE
: CITY: HILLSBOROUGH
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94010
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/172,528
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/937,466
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: OSMAN, RICHARD A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: T97-009
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 343-4341
: TELEFAX: (650) 343-4342
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1949 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-09-172-528-3

```

```

alignment_scores:
  Quality: 472.50      Length: 300

```


Ratio: 2.386 Gaps: 7
Percent Similarity: 66.000 Percent Identity: 35.667

alignment_block:
US-09-397-342-1 x US-09-172-528-3 ..

Align seg 1/1 to: US-09-172-528-3 from: 1 to: 1949

```

22 LysPheLeuLeuSerGlyCysAlaIaIaThrValAlaGluLeuAlaThrPh 38
|||||
285 AAGTTCCTGGGGGGCGGCGACCTGCTGTTTGGGACCTCTCCTCATT 334
|||||
38 eProLeuAspLeuThrThrArgLeuGlnMetGlnGlyLualaIaIaL 55
|||||
335 TCCCTGGACACCGCCAAAGTCCGCTGCATGCAAGGGAG..... 377
|||||
55 euaIaIaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMet 71
|||
378 .....AACCCAGGGCT...CAGAGCTGTCAGTACCGCGGTGTG 413
|||||
72 ValArgThrAlaLeuGlyIleIleGluGluGluGlyPheLeuLysLeuTr 88
|||||
414 CTGGGTACCACTGACTGATGTGCGACAGAGGGTCCCGCAGCCCTA 463
|||||
88 pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGlyA 105
|||||
464 CAGCGGACCTGGTCCGCTGCGCCACCGCCAGATGTTTGCCTTCATT 513
|||||
105 rGmetValThrTyrGlnHisLeuArgGluValAlaPheGlyLysSerGlu 121
|||||
514 GAATTGGCTCTAGACACTGTCTCAAGCAGTTCTACACCCCAAGGAGCG 563
|||||
122 AspGlnHisTyrProLeuTyrPlySerValIleGlyGlyMetMetAlaG1 138
|||||
564 GAC...CACCTCAGCGTCGCATCAGATCTGCGACGCTCGCACGACG 610
|||||
138 YValIleGlyGlnPheLeuAlaAsnProThrAspLeuValLysValGlnM 155
|||||
611 AGCCATGGCAGTGCACCTGCGCCACCGCCAGGATGTGTAAGGTCCGAT 660
|||||
155 etGln.....MetGlnGlyLysArgLysLeuGlnGly 165
|||||
661 TTCAAGCCATGATACGCTGGGAAGTGGAGAGAGAGAAA..... 701
|||||
166 LysProLeuArgPheArgGlyValHisIleAlaPheAlaLysIleLeuAl 182
|||||
702 .....TACAGAGGAGCTATGATGCTACAGACCATGCCAG 739
|||||
182 aGluGlyGlyIleArgGlyLeuTyrPalaGlyTyrPalaProAsnIleGlnA 199
|||||
740 GGAGGAAGGAGTCAAGGCTGTGGAAGGAGCTTGGCCCAACATCACAA 789
|||||
199 rGAlaAlaLeuValAsnMetGlyAspLeuThrThrTyrAspThrValLys 215
|||||
790 GAAATGCCATGTCAACTGTGCTGATGATGATGATGATGATGATGATG 839
|||||
216 HisTyrLeuValLeuAsnThrProLeuGlnAspAsnIleMetThrHisG1 232
|||||
840 GAGAACTTGGCTGAGCTTCACCTGTTACTGCAACACTTCCCTGTCTC 889
|||||
232 yLeuSerSerLeuGlySerGlyLeuValAlaSerIleLeuGlyThrProA 249
|||||
890 TGTCTCTCTTGGAGCTGGCTTGTGTCACACAGTGGTGGCTCCCGG 939
|||||
249 IAspValIleLysSerArgIleMetAsnGlnProArgAspLysGlnGly 265
|||||
940 TGGATGTGTAAGACCCGATCATGAAAGCTCC.....CTAGGC 980
|||||
266 ArgGlyLeuLeuTyrLysSerSerThrAspCysLeuIleGlnAlaValG1 282
|||||
981 AGG.....TACCGCAGCCCTCTGCATGATGATGATGATGATGATG 1021
|||||
282 nGlyGluGlyPheMetSerLeuTyrLysGlyPheLeuProSerThrPLeuA 299

```

```

|||||
1022 TCAGAGGAGGCCACGCGCTCTACAAAGATTTGTGCTTCCTTCTGC 1071
|||||
299 rGmetThrProTyrSerMetValPheThrPLeuThrTyrGlnLysIleArg 315
|||||
1072 GTCTGGAGCTTGGAACTGATGATGTTTGTACATATGACCAACTGAAG 1121

```

seq_name: /cgn2_6/prodata/1/lna/6A_COMB.seq:US-09-318-199-3

seq_documentation_block:

```

? Sequence 3, Application US/09318199
? Patent No. 6025469
? GENERAL INFORMATION:
? APPLICANT: Zhang, Ning
? APPLICANT: Amaral, M. Catherine
? APPLICANT: Chen, Jin-Long
? TITLE OF INVENTION: UCP3 Genes
? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
? STREET: 75 DENISE DRIVE
? CITY: HILLSBOROUGH
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 94010
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/318,199
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/937,466
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: OSMAN, RICHARD A
? REGISTRATION NUMBER: 36,627
? REFERENCE/DOCKET NUMBER: T97-009
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 343-4341
? TELEFAX: (650) 343-4342
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1949 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-09-318-199-3

```

alignment_scores:

Quality:	Ratio:	Length:
472.50	2.386	300
Percent Similarity: 66.000	Percent Identity: 35.667	Gaps: 7

alignment_block:
US-09-397-342-1 x US-09-318-199-3 ..

Align seg 1/1 to: US-09-318-199-3 from: 1 to: 1949

```

22 LysPheLeuLeuSerGlyCysAlaIaIaThrValAlaGluLeuAlaThrPh 38
|||||
285 AAGTTCCTGGGGGGCGGCGACCTGCTGTTTGGGACCTCTCCTCATT 334
|||||
38 eProLeuAspLeuThrThrArgLeuGlnMetGlnGlyLualaIaIaL 55
|||||
335 TCCCTGGACACCGCCAAAGTCCGCTGCATGCAAGGGAG..... 377
|||||
55 euaIaIaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMet 71

```

```

378 .....AACCCAGGGGCT...CAGAGCGTGCAGTACCCGCGGTG 413
72 ValArgThrAlaLeuGlyIleIleGlyGluGlyPheLeuYsLeuTr 88
414 CTGGGTACCATCTCTGACTATGTCGCGACAGAGGTCCCGGACCCCTTA 463
88 pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
464 CACGCGACTGCTGCGCTGCACCGCCAGATGATGTTGGCTCCATCC 513
105 rgmEtValThrTyrGlnHisLeuArgGlnValAlaPheGlyYsSerGlu 121
514 GAATGGCCCTTACGACTGTGTCAAGAGTTCTACACCCCAAGGGAGCG 563
122 AspGlnHisTyrProLeuTrPlySerValIleGlyGlyMetMetAlaG 138
564 GAC...CACGCCAGCGTCCGCATCAGATTTCTGCGAGCTGCACGACG 610
138 yValIleGlyGlnPheLeuAlaAsnProThrAspLeuValValGlnM 155
611 AGCCATGGCAGTGCAGCTGCGCCAGCCAGATGTGGTGAAGTCCGAT 660
155 etGln.....MetGlnGlyYsArgGlyYsLeuGly 165
661 TTCAGGCCATGATACGCTGGGAGACTGAGAGAGAGAGAA..... 701
166 LysProLeuArgPheArgGlyValHisHisAlaPheAlaYsIleLeuAl 182
702 .....TACAGAGGAGACTATGAGATCCCTACAGAACCATCCGCGAC 739
182 agGlyGlyGlyIleArgGlyLeuTrPalaGlyTyrPalaProAsnIleGln 199
740 GGAGGAGAGGAGTCAAGGCGCTGTGAAAGGAGCTTGGCCCAACATCAAA 789
199 rGAlaIleLeuValAsnMetGlyAspLeuThrThrTyrAspThrValYs 215
790 GAATGCGCATTTGCAACTGTGTGAGATGGTGACTACGACATCATCAAG 839
216 HisTyrLeuValLeuAsnThrProLeuGluAspAsnIleMetThrHisG 232
840 GAGAGTGTGCTGAGTCTCACTGTTTACGACACATCCCGCTGACTT 889
232 yLeuSerSerLeuCySerGlyLeuValAlaSerIleLeuGlyThrPro 249
890 TGTCTGTGCTTGGAGCTGGCTTCTGTCCACAGATGGTGGCTCCCGG 939
249 IaAspValIleYsSerArgIleMetAsnGlnProArgAspYsGlnGly 265
940 TGGATGTGGTAAAGACCCGATACATGACCTCC.....CTAGGC 980
266 ArgGlyLeuLeuTyrLysSerSerThrAspCyLeuIleGlnAlaValG 282
981 AGG.....TACGCGAGCCCTCGCACTGTATGCTGANAATGGTGGC 1021
282 nGlyGlyGlyPheMetSerLeuTyrYsGlyPheLeuProSerTrPlea 299
1022 TCAGGAGGAGCCACCGCCTTCTACAAAGATTTGCCCCCTTCTGCG 1071
299 rgmEtValThrTyrPseMetValPheTrPheThrTyrGlnYsIleArg 315
1072 GTCTGGGAGCTTGGACGTCGATGATGTTGTAAACATATGAGCAAGTGA 1121

```

```

seq_name: /cgn2_6/ptodata/1/1na/5B_COMB.seq:US-08-937-466-1
seq_documentation_block:
; Sequence 1, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes

```

```

? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
? STREET: 75 DENISE DRIVE
? CITY: HILLSBOROUGH
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 94010
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/937,466
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: OSMAN, RICHARD A
? REGISTRATION NUMBER: 36,627
? REFERENCE/DOCKET NUMBER: T97-009
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 343-4341
? TELEFAX: (650) 343-4342
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2782 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-08-937-466-1

```

```

alignment_scores:
Quality: 469.50 Length: 314
Ratio: 2.313 Gaps: 8
Percent Similarity: 64.650 Percent Identity: 34.713

```

```

alignment_block:
US-09-397-342-1 x US-08-937-466-1 ..

```

```

Align seg 1/1 to: US-08-937-466-1 from: 1 to: 2782

```

```

22 LysPheLeuLeuSerGlyCysAlaAlaThrValAlaGluLeuAlaThrP 38
285 AACTTCTGCGGGGCGCGCACTGCGGCTGTTTGGCGACTCTCTACTT 334
38 eProLeuAspLeuThrLysThrArgLeuGlnMetGlnGlyGlnAlaAla 55
335 TCCCTGAGACACCGCCAAAGGTCCGTCTGCAAGATCCAAAGGGAG..... 377
55 euAlaArgLeuGlyAspGlyAlaArgGluSerAlaProTyrArgGlyMet 71
378 .....AACCCAGGGGCT...CAGAGCGTGCAGTACCCGCGGTG 413
72 ValArgThrAlaLeuGlyIleIleGlyGluGlyPheLeuYsLeuTr 88
414 CTGGGTACCATCTCTGACTATGTCGCGACAGAGGTCCCGGACCCCTTA 463
88 pGlnGlyValThrProAlaIleTyrArgHisValValTyrSerGlyGly 105
464 CACGCGACTGCTGCGCTGCACCGCCAGATGATGTTGGCTCCATCC 513
105 rgmEtValThrTyrGlnHisLeuArgGlnValAlaPheGlyYsSerGlu 121
514 GAATGGCCCTTACGACTGTGTCAAGAGTTCTACACCCCAAGGGAGCG 563
122 AspGlnHisTyrProLeuTrPlySerValIleGlyGlyMetMetAlaG 138
564 GAC...CACGCCAGCGTCCGCATCAGATTTCTGCGAGCTGCACGACG 610
138 yValIleGlyGlnPheLeuAlaAsnProThrAspLeuValValGlnM 155

```

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,466
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SRO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2782 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-172-528-1

alignment_scores:
Quality: 469.50 Length: 314
Ratio: 2.313 Gaps: 8
Percent Similarity: 64.650 Percent Identity: 34.713

alignment_block:
US-09-397-342-1 x US-09-172-528-1 ..

Align seg 1/1 to: US-09-172-528-1 from: 1 to: 2782

22 LysPheLeuSerGlyCysAlaIleThrValAlaGluLeuAlaIleThrPhe 38
||||| :||| :||| | ||| :|||
285 AGTTCTCGGGGGCCGGACAGTGGCGCTGTTCGGGACCTCCACTTT 334
||||| :||| :||| :||| :|||
38 eProLeuAspLeuThrLysThrArgLeuGlnMetGlnGlyAlaIleAla 55
||||| :||| :||| :||| :|||
335 TCSCCTGGCACCCGCCAAGAGTGCTCTGCATCCCAAGGGAG..... 377
||||| :||| :||| :||| :|||
55 euAlaArgLeuGlnLysrPheIleAlaHrgIleSerAlaIleProTyArgGlyMet 71
||||| :||| :||| :||| :|||
378 .....AACCCAGGGGCT...CAGACGTGCACATACCAGCGGGTGG 413
||||| :||| :||| :||| :|||
72 ValArgThrAlaLeuGlnIleIleGlnGlnGlnLysPheLeuLysLeuTr 88
::: ||| :||| :||| :||| :|||
414 CCGGGTACCACATCTCATATGATGGTGGGCACAGAGGGTCCCCGCCACCCCTA 465
||||| :||| :||| :||| :|||
88 rGlnGlnValThrProAlaIleTyArgHisValValTyLysSerGlyLys 105
::::: ||||| :||| :||| :||| :|||
464 CAGCGGACGTGGTGGCTGGGCTGCACACGCCACAGATGATTTGCTCCATTC 513
||||| :||| :||| :||| :|||
105 rgMeValThrTyTrcIleHisLeuAlaHrgIleValAlaPheGlnLysSerGln 121
||||| :||| :||| :||| :|||
514 GAATTGGCGCTACAGACCTGTGCAGAAGCAAGTTCTACACCCCACAAGGGAGCG 563
||||| :||| :||| :||| :|||
122 AsgGlnHisTyTrcProLeuThrLysSerValIleGlnLysMetIleAlaG 138
||| ||| :||| :||| :||| :|||
564 GAC...CACTCAGCGGTGCCCATCAGAGATTTGTGGCAGGCTGCCACAGACAG 610
||||| :||| :||| :||| :|||
138 yValIleGlnPheLeuAlaAsnProThrAspLeuValLysValGlnM 155
||||| :||| :||| :||| :|||
611 ACCCATGGCAGTAGACCTCGCCACAGCCACACAGATGTGGTGAAGGTCCGAT 666
||||| :||| :||| :||| :|||
155 etcGln.....MetGlnGlnLysrArgLysLeuGlnGly 165
||||| :||| :||| :||| :|||
661 TTCAAAGCATGATACGCTGGGAACTGGAGAGAGAGAAA..... 701
||||| :||| :||| :||| :|||
166 LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIleLeuAl 182
||||| :||| :||| :||| :|||
702 .....TACAGAGGACATGAGATGCCATACAGAACATCCGCAG 739
||||| :||| :||| :||| :|||
182 agIuGlnGlyIleArgGlyLeuTyrrAlaGlyTYrrValProAsnIleGlnA 199
||||| :||| :||| :||| :|||

```

```

740 GGAGGAGAGAGTCAGGCGCTGTGGAAGGAGCTGGCCCAACATCAACA 789
199 rgAlaIaIeUaValaSnMeTgLYaSPLeuThrThyAspThrValys 215
790 GAATATGCAATGTCAACTGTGTGAGATGTGTGACCTACGACATCATCA 839
216 HlSTyLeuValLeuAsnThrProLeuGluAspAsnIlleMethrHlsgl 232
840 GAGAACTGTGCTGAGTCTCACTGTTACTGACACACTTCCCTGTCACTT 889
232 yLeuSerSerLeuCySerGlyLeuValAlaSerIlleLeuGlyThPro 249
890 TGTCTGTGCTTGGAGCTGTGTCTGTGCAACAGTGTGTGCTCCCTCC 939
249 lAspValIlleLysSerArgIlleMetAsnGlnProArgAspLysGln 265
940 TGGATGTGTGAAGACCCGATACATGACGCTCC.....CAGAGC 980
266 ArgGlyLeuLeuTyTrLysSerSerThrAspCysLeuIlleGlnAlaVal 282
981 AGG.....TACCGAGCCCTGTGCACTGTATGCTGAGAGATGGCGG 1021
282 ngLYgluGlyPheMetSerLeuTyTrLysGlyPheLeuProSerTrpLeu 299
1022 TCAGGAGGAGCCACGCGCTTCTACAAAGGATTGTGCTCTCTCTCTGC 1071
299 rgMethrTrpTrpSerMetValPheTrpLeuThyTrpGluLysIleArg 315
1072 GTCTGGAGAGCTTGAACGTGATGTTGTCTTACATATGACACATGAG 1121
316 .....GlMetSerGlyValSerProPhe 323
1122 AGGCGCTTAATGAAGTCCAGTACGCGGGAATCTCCGTTT 1163
seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-318-199-1
seq_documentation_block:
; Sequence 1, Application US/09318199
; Patent No. 6025469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/318.199
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 2782 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-09-318-199-1

alignment_scores:
    Quality: 469.50      Length: 314
    Ratio: 2.313        Gaps: 8
    Percent Similarity: 64.650    Percent Identity: 34.713

alignment_block:
US-09-397-342-1 x US-09-318-199-1

Align seg 1/1 to: US-09-318-199-1 from: 1 to: 2782

22 LysPheLeuLeuSerGlyCysAlaIaIaThrValAlaGluLeuAlaThrP 38
285 AAGTTCCTGGGGCGCGGCACTGCGGCTGTTTGGCGAGCTCTCACTT 334
38 eProlaAspLeuThrLysThrArgLeuGlnMetGlnGlyAlaIaIa 55
335 TCCCTGGACACCGCCAGGTCGCTGCAAGATCAAGGGAG..... 377
55 euAlaArgLeuGlyAspGlyAlaArgGluSerAlaProTrpArgGlyMet 71
378 .....AACCCAGGGGCT..CAGAGCTGCAGTACCGCGGTGTC 413
72 ValArgThraIaLeuGlyIleIlleGluGluGlyPheLeuLysLeuTr 88
414 CTGGGTACCACTCTGACTATGTCGCGACAGAGGTGCCGCGAGCCCTTA 463
88 pgLingValaThrProAlaIleTrpArgHisValaValaTySerGly 105
464 CACGGAGCTGTGCTGCTGCTGCACCGCCAGATGATTGCTCCATTC 513
105 rgMethrThrTrpGlnHisLeuArgGluValaValaPheGlyLysSer 121
514 GAATTGGCCCTTACGACTGTGTCAAGCAGTTTACACCCCAAGGAGCG 563
122 AspGlnHisTrpProLeuTrpLysSerValIlleGlyMetAlaG 138
564 GAC...CACTCCAGCGCGCCATCAGAGATTCTGCGAGGCTGCACGAC 610
138 yValIlleGlyGlnPheLeuAlaAsnProThrAspLeuValaGlnM 155
611 AGCCATGGCAGTGCCTGCGCCAGCCACGAGTGTGTGAAGTCCGAT 660
155 etGln.....MetGluGlyLysArgLysLeuGluGly 165
661 TTCAGCCATGATACCGCTGGAACTGGAGAGAGAGAAA..... 701
166 LysProLeuArgPheArgGlyValHisHisAlaPheAlaLysIlleu 182
702 .....TACAGAGGACTATGATGATGCTACAGAAACATGCCAG 739
182 agLyuGlyLysIleArgGlyLeuTrpAlaGlyTrpValProAsnIlle 199
740 GGAGGAGAGAGTCAGGCGCTGTGGAAGGAGCTGGCCCAACATCAACA 789
199 rgAlaIaIeUaValaSnMeTgLYaSPLeuThrThyAspThrValys 215
790 GAAATGCAATGTCAACTGTGTGAGATGTGTGACCTACGACATCATCA 839
216 HlSTyLeuValLeuAsnThrProLeuGluAspAsnIlleMethrHlsgl 232
840 GAAATGCTGTGAGTCTCACTGTTACTGACACACTTCCCTGTCACTT 889
232 yLeuSerSerLeuCySerGlyLeuValAlaSerIlleLeuGlyThPro 249
890 TGTCTGTGCTTGGAGCTGTGTCTGTGCAACAGTGTGTGCTCCCTCC 939

```

```

249 1AASPVALIILEYSSERARGILEMETASNGINPROARGASPLYSGINGLY 265
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
940 TGGATGTGGTAAGACCGATACATGACGCTCC.....CTAGGC 980
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 ARGGLYLEUPTYLYSSERSETRHASPYSLEULEGIALAVALGI 282
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
981 AGG.....TACCGACGCCCTCTGCACGTGATGCGAGAGATGGCGGC 1021
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 NGLYGLUPLYHMETSERLEUTYLYSGLYPHELEUPROSETRIPLEUA 299
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1022 TCAGAGGAGGACCCAGCGCTCTCTACAAAGATTTGCGCTCTTCGCGC 1071
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 RGMETHPROTPRSEMETVALPHETRPLEUTHTYRGLULYSLEIARG 315
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1072 GTCTGGAGCGCTTGGACGTGATGATGTTGTACATATGAGCACTGAG 1121
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 .....GlumetserglyValSerPrope 323
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1122 AGGCGCTTAATGAAGTCAGGTACTGGGGAATCTCCGTTT 1163
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: /cgn2_6/prodata/1/ina/5A_COMB.seq:US-08-518-878B-36

seq_documentation_block:
: Sequence 36, Application US/0851878B
: Patent No. 5702902
: GENERAL INFORMATION:
: APPLICANT: Tartaglia, Louis A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
: TITILE OF INVENTION: TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/518, 878B
: FILING DATE: 23-AUG-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7853-036
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 36:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1205 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-518-878B-36

alignment_scores:
: Quality: 396.50      Length: 316
: Ratio: 1.925        Gaps: 12
: Percent Similarity: 65.190      Percent Identity: 33.228

alignment_block:
US-09-397-342-1 x US-08-518-878B-36 ..
Align seg 1/1 to: US-08-518-878B-36 from: 1 to: 1205

```

```

18 PROARGALASER...LYSPHELEULEUSERGLYCYSALAALATHRVALL 33
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
31 CCACAGCCACTGTGATGATTCCTGGGGCTGGACAGCTGCTGCATTCG 80
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 AGLULEUALATHRPHEPROLEUASPLEUTHRLYSTRARGLEUNGIMETG 50
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81 AGATCTCATCACTTCCCTCTGAGATACGCCAAGSTCGCGGTGAGATCC 130
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 INGLYGLUALA...ALALEUALARGLEUNGLYASPLYALARGLUSER 65
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
131 AAGGGAGAGTCAGAGGCTAGTGCCTC.....ACCGCAGCCAGC 168
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 ALAPROTYRARGLYMETVALARGHRALEUNGILEIIEGLINGLUGI 82
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 GCCCATCCGGGGGCTGTCTGGTACCAATCCCAACCAATGGCGCATGA 218
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
82 UGLYPHELEUPLYSEUTRPGINGLYVALTHRPOLALLETYRARGHSV 99
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 GGGTCACAGCAGCCCTCTCAATAGGCGTGGTGGCGCGCCTGCAGCGC 268
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
99 ALVALTYRSEGLYGLYARGMETVALTHRYRGLUHSLEUARGLUVAL 115
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 TGAGCCTTGCCCTCCGCTCCGATGGCCTCTACGACTCTGCAMACAG... 315
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 VALPHEGLYLYSERGLUASPOLUHSITRYRPROLEUTRPLYSERVALI 132
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 ...TTCTACACCAAGGCTCCAGAGCATGAGGACATGGGAGCCGCTCT 362
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 EGLYGLYMETMETALAGLYVALIIEGLYLPHELEUALAASPROTHRA 149
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 GGCAGGTAGCACCAACAGAGGCGCTGGGTGTACCCAGCCTACAG 412
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 SPLEUVALLYSVALGIMETGINMETGLUGLYLSARGLYSLEUNGILUG 165
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 ATGTGTAAAGTCCGCTTCCAGGCTCCAGCGCG.....CGTGT 453
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 LYSPROLEUARGPHEARGLYVALHISHIALPHEALILYILEUAL 182
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 GGTGCTCGAGATACAGACACTG.....TCGACCTCAACAAACATCAC 497
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 AGLUGLYLYLEARGLYLEUTRPLAGLYTRVALPROASNILEGNA 199
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
498 GAGAGAGAGGAGATCCGGGCGCTCTGGAAGGA...CTGCCAATGTGCC 544
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 RGALALALEUVALASMETGLYASPLEUTHRLTYRASPTHVALYS 215
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
545 GTAATGCCATGTGTCAACTGTGCTGAGCTGACCTATGACCTCATCAAA 594
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 HISTYRLEUVALLEUASNTHRPROLEUGLUASPASNILEMETHHSGL 232
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
595 GATACTCTC...CTGAGCCACTCATGACAGATGACCTCCCTGCACACT 641
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 YLEUSERSERLEUCYSERGLYLEUVALALASERILLEUNGILYTHRPROA 249
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
642 CACTTCTGCTCGGGGGGGGCTCTGCACCAACCTCATCGGCTCCCTCG 691
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 1AASPVALIILEYSSERARGILEMETASNGINPROARGASPLYSGINGLY 265
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
692 TGGATGTGGTCAAGACGAGATACATGACT..... 720
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 ARGGLYLEU...TYLYSSERSETRHASPYSLEULEGIAL 280
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721 .....CTGCTGGCCAGTACCAAGCGAGTCACTGTGCCCTTACATG 764
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 AVALGINGLYGLUPLYHMETSERLEUTYLYSGLYPHELEUPROSETR 297
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
765 CTCGGAG...GAGGAGCCGCGCTCTCAACAGAGGGGTATAGCCTTCT 811
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 RPLEUARGMETHRPROTPRSEMETVALPHETRPLEUTHRYRGLULYS 313
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
812 TTCTCCTGTGGATCTGTGAACGTAGTGTGTTGTACCATATGAGCAG 861
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 16:17:21 ; Search time 1369.25 Seconds
(without alignments)
11190.817 Million cell updates/sec

Title: US-09-397-342-2
Perfect score: 1039
Sequence: 1 ccgagctcgagcccgctat.....cagatatacctacactgac 1039

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues
Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_rod: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v1: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rol: *
95: gb_rol2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	972	93.6	972	88	AF110532
2	785	75.6	2678	95	RNO300162
3	716.8	69.0	1890	95	RNO300164
4	715.2	68.8	2620	95	RNO300163
5	676.8	65.1	1726	9	AX061217
6	194.6	18.7	55820	81	AL512483
7	194.6	18.7	185614	60	AC008104
8	134	12.9	185614	60	AC008104
9	124	11.9	112126	12	AC007576
10	117.4	11.3	1525	12	AB024733
11	116.2	11.2	1434	15	STUCPMRNA
					Y11220 Solanum tub

12	116	11.2	169457	61	AC0092162	AC0092162	Drosophila
13	116	11.2	191504	63	AC0021162	AC0021162	Drosophila
14	116	11.2	300994	4	AE003506	AE003506	Drosophila
15	115	11.1	1240	13	ATMUNUCU	ATMUNUCU	Drosophila
16	114.4	11.0	79977	63	AC012693	AC012693	Drosophila
17	113.4	10.9	921	12	AE146226	AE146226	Arabidopsis
18	113.4	10.9	1225	13	ATPUMP	AT1229983	Arabidopsis
19	112.8	10.9	328	54	G22138	G22138	human
20	104.6	10.1	1088	12	AB021706	AB021706	human
21	100.6	9.7	140468	65	AC017377	AC017377	Drosophila
22	100.6	9.7	174311	4	AC004758	AC004758	Drosophila
23	100.6	9.7	260550	4	AE003612	AE003612	Drosophila
24	100.4	9.7	18557	5	CEKK07B1	AF003384	Caenorhab
25	90.6	8.7	2991	12	AB024734	AB024734	Symlocas
26	88	8.5	969	94	AF155813	AF155813	Mus muscu
27	88	8.5	975	94	AF155812	AF155812	Mus muscu
28	88	8.5	1775	94	AF076981	AF076981	Mus muscu
29	87.2	8.4	971	12	AB044428	AB044428	Triticum
30	85.6	8.2	974	12	AB042429	AB042429	Triticum
31	80	7.7	969	94	AB026879	AB026879	Rattus no
32	80	7.7	997	95	RNO300165	AF0300165	Rattus no
33	80	7.7	1572	94	AF300424	AF300424	Rattus no
34	78.4	7.5	969	88	AF155811	AF155811	Homo sapi
35	78.4	7.5	978	88	AF155809	AF155809	Homo sapi
36	78.4	7.5	1652	88	AF078544	AF078544	Homo sapi
37	77.8	7.5	125803	13	ATF22K18	AF035356	Arabidops
38	77.8	7.5	198402	13	ATFCHT161	AF1615161	Arabidops
39	74.2	7.1	102051	12	AC006592	AC006592	Arabidops
40	74	7.1	5119	94	AB011068	AB011068	Rattus re
41	73.2	7.0	1297	14	NTA29250	AJ299250	Nicotiana
42	60.8	5.9	1398	9	AX006066	AX006066	Sequence
43	60.8	5.9	1398	15	STOKMATR	ST98853	S. tiberdun
44	58.8	5.7	1313	7	RNUCEP24	X14064	Bovine mRNA
45	57.2	5.5	924	9	AX062004	AX062004	Sequence

ALIGNMENTS

AC009216 Drosophila
AC012162 Drosophila
AE003506 Drosophila
AJ001124 Arabidopsis
AC012653 Drosophila
AF116226 Arabidopsis
AF233963 Arabidopsis
G22138 human SRS Y
AB021706 Arabidopsis
AC011737 Drosophila
AC003718 Drosophila
AE004652 Drosophila
AF003362 Drosophila
AF003434 Caenorhab
AB024774 Symplodia
AF155813 Mus musc
AF155812 Mus musc
AF076981 Mus musc
AB042429 Trillium
AB042428 Trillium
AB028879 Rattus no
AF300145 Rattus no
AF300144 Rattus no
AF155811 Homo sapi
AF155809 Homo sapi
AF078544 Homo sapi
AL161551 Arabidopsis
AL161551 Arabidopsis
AC006552 Arabidopsis
AB011068 Rattus r
AJ292920 Nicotiana
AX006066 Sequence
X98953 S. tiberosun
X14064 Bovine mRNA
X5062004 Sequence

```

RESULT      1
AF110532
LOCUS
DEFINITION Homo sapiens uncoupling protein UCP-4 mRNA, nuclear gene encoding
              mitochondrial protein, complete cds.
ACCESSION
VERSION     AF110532.1
KEYWORDS    GI:4324700
SOURCE
            human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 972)
            Mao, W., Yu, X.X., Zhong, A., Li, W., Brush, J., Sherwood, S.W.,
            Adams, S.H. and Pan, G.
            UCP4, a novel brain-specific mitochondrial protein that reduces
            membrane potential in mammalian cells
            FEBS Lett. 443 (3), 326-330 (1999)
            99148824
AUTHORS     2 (bases 1 to 972)
            Pan, G., Mao, W., Zhong, A. and Brush, J.
            Direct Submission
            Submitted (02-DEC-1998) Endocrinology, Genentech, Inc., 1 DNA Way,
            M/S-37, South San Francisco, CA 94080, USA
            location/Qualifiers
FEATURES
            source
            1..972
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p11.2-q12"
            /tissue_type="brain"
            1..972
            /function="uncouples respiration in mammalian cells"
            /codon_start=1

```

Query Match	Best Local Similarity	Score 972:	DB 88:	Length 972:
Matches 972:	Conservative	0:	Mismatches	0:
Indels	0:	Gaps	0:	
0y	40	atgtccgcgccggaggaaggaagaggttttgcgcgtgacccagagaatgccccgagcg	99	
Db	1	ATGTCCGTCGCCGAGAGAGAGAGGCGCTTTTGCCGCTGACCCAGAGATGGCCCCGAGCG	60	
0y	100	agcaaatcttaccatgctccgcgtgctgcgcgtctacgcgtgctgcgcgtctagcaacttccctg	159	
Db	61	AGCAAAATTCCTAAGTCTCCGGCTGCCGGCTACCGTGGCCGAGCTAGCAACTTTCCTTG	120	
0y	160	gattcacaaaaactcgactcccaatgcaagagaagcagctcttgcctcgttctgggaagac	219	
Db	121	GATCTCACAAAACCTCGACTCCAAATGCAAGAGACAGCTTGTGCTCGGTTGGAGAC	180	
0y	220	ggtgcagaagaatctgcgcctctatagggaatggttgcgcagaagccctaggatcatatgaa	279	
Db	181	GGTGCAAGAGAAATCGCCCCCTATAGGGGAATGGTGGCCACAGCCCTTAGGGATCAATTGAA	240	
0y	280	gaggaaggcttcttcaagcttctggcaaggaatgcaaccgcgcatttacaagacagatgctg	339	
Db	241	GAGGAAGCCTTCTTAAGCTTTGGCAAGAGAGTGAACCCGCCATTTACAGACAGTATGTG	300	
0y	340	tattctggaggtcgaaatggtcacaatalgacaatctcgaagaggttggtttggcaaaagt	399	
Db	301	TATTTGTGAGGTCGCAATGTCACATATGACATCTCCAGAGGTTGTGTGGTGGCAAAAGT	360	
0y	400	gaaagatgagcatatccccccttggaaatcgaatcttggaaagatgaaatgagctgcggttatt	459	
Db	361	GAAGATGAGCATATATCCCTTTGGAAATCAGTCATGTGAAGGAGGATGATGGCTGTGTATTT	420	
0y	460	ggccaggtttttagccaatcccaactgaacctagtgaaagcttccagatgcaaatgaaagaaaa	519	
Db	421	GGCCAGTTTTAAGCCAAATCCAACTGACTTGTGAAGTTCAAGTTCGAATGGAAGGAAAA	480	
0y	520	aggaactctgaaagaaaaccatctgcgattcttgcgttgatgatacatgcaatcttgcaaaaatc	579	
Db	481	AGGAACATGGAAGAAACCAATTTGGGATTTGCTGTGTATCATGTCATGTTGCAAAAAATC	540	
0y	580	tttagctgaaggaagaatactcgaaggtcttggcagagctgcyggtaccacaatacaaaagca	639	
Db	541	TTTAGCTGAAGAGCAATATCGAGGCGCTTTGGGAGGCTGGTGAACCAAAATTCAAAAGACA	600	
0y	640	gcaactgctggaatatgagagatttaacccactatgatacagtgaaagaaactacttgatttg	699	
Db	601	GCACTGGTGAATATGGGAGATTTTAACCACTTATGATACAGTGAAGAACACTACTTGTGATTG	660	
0y	700	aatacaacaacttggagaaatatacatgactcaggttcatataagttttagtcttgcgagct	759	
Db	661	AATNACACACTTGAAGAGCAATATCATGACTGACGCTTATCAAGTTTATGTTTCGAGCTG	720	
0y	760	gtagctctatctctgggaacaccagcogaatgltcaaaagacagaataatgaatcaacca	819	
Db	721	GTAGCTTCTATTTCTGGGAACACCAAGCCATGTCTATCAAAAACAGAAATATGAATCAACA	780	
0y	820	cgagataaacaagaagggagacttttgataataatcatgactgactgctgttattccagct	879	
Db	781	CGAGATAAACAAGGAAGGGAGCTTTGTATTAATCATGACTGACTGCTGTATTCAGGCT	840	

Db	32	TGTTCTTACACCTGGTGTGTAATGCTTTTCCC---TGAGGAGGAGTTCGGTACAGGCCACTTCA	88
QY	80	cccaagatctggccccagcgagcaaatctcctaactctccgctctgcggtctacgcttgccg	139
Db	89	CGCAAGATATGGCCCCCGGACACCAAGTTCTCTACTCTCCGGCTGCCGCCCAACCTTGCGCG	148
QY	140	agcttgagcaactcttcccccttgatctcacaanaactcgcgtcccaatctgaagaagaacag	199
Db	149	AGCTAGCAACCTTTTCCCTCGATCTCACAAAAACCCGACTCCAGATGCAAGGAGAACTGG	208
QY	200	ctctgtctcgtcttggaagacgctgaagaagaatctgcgccccatctagggaacgctgcga	259
Db	209	CCCTTGCTACTAAGTTGGGAGATGCTGGCATGGAATCCGCCCTTACAGGGGCATGATGCGCA	268
QY	260	cagcccttaggagatctgaagaagaaggctcttctaaggctcttgcaagagatgcacccg	319
Db	269	CGGCCCTGGGAGTTTGTCCAGAGGAAAGGTTCTCTAAAGTTGAGCAAGAGTGCAGCCCG	328
QY	320	ccatttacaagacgtaagtatctcttgaaagctgcgaatgttcacatatgaacatctccgag	379
Db	329	CCATTTACAGACACTAGTGTACTCTGGAGGTCGGATGTGCTACCTACGAMCATCTCCGG	388
QY	380	aggtgtctcttgcaaaaagtgaagaatgcataatccccctttggaaatcaatgcacttgag	439
Db	389	AAGTTGTGTTTGCCAAAAGTGAAGATGACATTAATCCCTCTGGAAATGCCGATTTGGAG	448
QY	440	ggaatgatgcctgctgtatcttgccagcttcttagccaatccactagctagtgaaagtctc	499
Db	449	GSATATATGGCTGGCTCATCTACGACAGTTTTATGCCAAOCCCACTACCTCGTGAAGTTC	508
QY	500	agatgcacaatggaagaaanaaggaacatggaagaaacacatctgcgattcgttggttac	559
Db	509	AGATCAAAATGGAAGGAAACCTTAACATGGAAGGAAACCCCTTGGAATCCCTGGAGTAC	568
QY	560	atcatgctcttgcaaaaactcttagcttgaaaggagaaatcgaaggtcttgggaagcttgg	619
Db	569	ATCATGCGTTTGCCAAAATCTTAGCTGAAGGAGAAATCGTGCGCTTGGSCAGGCTGGA	628
QY	620	taccacaataacaagagcagcactcgtgaatatatgaggagatctaaccacttaatgaacag	679
Db	629	TACCCAATATTTGAGAGGAGGACACTTTGSAACATGGGAGATTTAAOCACTTACGACACAG	688
QY	680	tgaacactacttgatlattgaatacaccaacttgaaggacaatatcatgactcaagcttat	739
Db	689	TGAACCACTACTCGGTATTTGAACACACGCACTTGAAGACAAATATCGCACTACGCGCTTAT	748
QY	740	caagtttatgtcttgagactcgaagctcttatctctcttggaacaccagccgagtgatcaaaa	799
Db	749	CCAGTTTGTGTTCTTGAGCTGGTGGCTCTTATTTCTGGGAACACCAAGCGGATGATCAAAA	808
QY	800	gcagaaataatgaatacaaccacagagataaacaagaagaaggactcttgataaataatcga	859
Db	809	GCGGAATATATGAAACCAACCTCGACGACAAACAAGAAAGAGCGTTTGTATTAAGTCAATCA	868
QY	860	cttactcttgatctcaagctgttcaagtgtaaggatctcaatgactataataaagcttt	919
Db	869	CTGACTGGGTGATTCAGAGCTGTTTCAAGGAGAGAGGCTTCTTGAGCCGTGTTAAAGCGCTTC	928
QY	920	taccatcttgcttgatgaatagaccctctgttcaatggtgtgtctgcgtctactatgaaaaa	979
Db	929	TATCCTCTCGCTGGTGAAAGACCCCTTGCTCAATGGTGTCTTGCGTTACTTATGAAAAAA	988
QY	980	tcaagagatgtagtgagatcagttcatctttaag 1012	
Db	989	TCAGACAGCTGAGTGGAGTCAGTCCATTCTAAG 1021	

RESULT

3

RNO300164

LOCUS

DEFINITION

Rattus norvegicus mRNA for uncoupling protein UCP-4 (Ucp-4 gene)

ACCESSION

AJ300164

1890 bp

RNA

06-JAN-2001

ROD

BASE COUNT	ORIGIN
657 a	661 c 679 g 623 t
	<pre> /strain="Sprague Dawley" /db_xref="taxon:10116" /sex="male" /tissue_type="brain, preoptic anterior hypothalamus" /dev_stage="6 weeks" /tissue_lib="UNI-ZAP custom cDNA library" /country="Switzerland" 139..1173 /gene="Ucp-4" 139..1173 /gene="Ucp-4" /function="uncouples respiration in eukaryotic cells" /codon_start=1 /product="uncoupling protein UCP-4, isoform b" /protein_id="CAC20899.1" /db_xref="GI:12055542" /translation="MPPEEESIQPIORMPRTSKFLISGCAATVATLAPPLDLITK RLQNGEALALAKLDGAMESAPRYGMMRTALGIVOEESFLKINGVTPAIRHYVYSG GRWTPYELHRELVFVHAKSDEDEYPLMKSVIGGMAGVIGQFLANPTDLVKVOMQEGKR RLEKTPPLRFGRVHAFKILAEGGIRIGMAGWIPINQALVANGDGLTYDTVHYLYLV LNTPLAEINATHGSLSCSGIVASIIILGPAIVISRLINQNRDQSGKGLYKSSDCV IOAVQEGFLSYLKGFLPSWLMEDTLLHRCDDLGATIVPSLETGLVSSSFYLRADH CLVOSHETNH" </pre>

	Query Match	68.8%	Score 715.2;	DB 95;	Length 2620;
	Best Local Similarity	86.8%;	Pred. No. 2,4e-192;		
	Matches 799;	Conservative 0;	Mismatches 118;	Indels 3;	Gaps 1;
OY	20 tgcgttcgctacatcgctcgaattcctccgccggagagagagagagagcttttgcgcgtca	79			
Dd	119 TGGTCTTGACCGTCGCTGTGAATGCCCTTTCCC---TGAGGAGGAGTGCCTACAGCCACTICA	175			
OY	80 ccagaagatlygccccagcgagcaaatctctaactlctgcgctctgcgcgctaacctlygcg	139			
Dd	176 GCGAAGAATGCGCCCCGCGACACACAAGTTCTCCTACTGTCGCGTCCGCGCCGCAACCTGTGCGCG	235			
OY	140 agctgcgaacctttccccttgatatcccaaaaaaccctcgactccaatltgaagaagaagcaag	199			
Dd	236 AGCTGACCAACCTTTCCTCCCTCGATCTCACAAAAAACCCGACTCCAGATGGAAGAAGAAGTGTG	295			
OY	200 ctctgcctcgtcttgaggagacggtlgaagaagaatctgcgcccatatagggaaatlytgcga	259			
Dd	296 CCGCTTGCTAAAGTGGTGGAGATGCTGGCATGTGAATCCGCCCTTACAGGGGCAATGATGCGCA	355			
OY	260 cagccctagggatcatlgaagaggaagccttctaagacttlygcaagagtgcacaccg	319			
Dd	356 CGGCGCTGGGGATTCTCCAGAGGAAGGCTTCCTAAAGTGGGCAAGAAGTAGACGCCCG	415			
OY	320 ccatttaagaagaacgtagtgtatctctgtagagtgcgaatygtacaataatcctccgag	379			
Dd	416 CCATTTAAGACACACTACTAGTCTCTTGAGAGCGGATGGTACCCTACACAACTTCCTCCGGG	475			
OY	380 aggtctgtcttgaccaaaagtgaatgaatgcattaccctcttggaatcaatgcattctgag	439			
Dd	476 AAGTTGTGTTTTGGCAAAGCTGAAGATGACANTATATCCCTCTGGAAAGTCCGTCATTGGAG	535			
OY	440 ggaatgatlygctgtgtltaattlgcacagtltttagccaatccaactlgaactagtgaagtgc	499			
Dd	536 GGATGATGGCTGGCGTCATCGCAGAGTTTTGAAGCCAAOCCTGACCTCGTGAAGGTCC	595			
OY	500 agatgcaaatgagaagaaaagaagaactcrggaaggaaaaacacatlcgatttgtgtgttac	559			
Dd	596 AGATCCAATAATGGAAAGGAAAACGTAACTCTGGAAGGGAAGAACCTTTGAGATTTCGCGGAGTAC	655			
OY	560 atcagaatcttgcaaaatcttagcttgaagaaggaatacgaaggtcttggagcaggtcg	619			
Dd	656 ATCATGCGGTTTCCAAAAATCTTAGCTGAAGAGGGAATCCGTGGCGCTTTGGCAGGCGTGA	715			
OY	620 taaccaataataaagaagcagcactlytgaatatalyggaatltaaccactiatatgaataag	679			

Db	716	TACCAATATATCAAGAGCAGCAGCACTTGTGACATGTGGAGATTTAACCACTTAGACACAG	775
QY	660	tgaaacactacttgytttttgaataacaccacttggaggaacaatatactagctaagttlat	739
Db	776	TGAAGCATTACTACTGGTATTTAAACAACGACCTTGAACATATTCGCCACTTACCGGCTTAT	835
QY	740	caagtttatgtctcgtagctgttagcttcatctatcttggagacaccagccgaltgtacaa	799
Db	836	CGAGTTGTGTGTTGGACTGCTGGCTTTATTTCTGAGGAACACAGCGATGTCATCAAA	895
QY	800	gcagaaataatgaatcaaccaccgcgaataaacaaggaaaggagacttttgtataataatcatcga	859
Db	896	GCCGATATATGAACCAACCTCGAGAACAAACAAGGAAGAGGCGCTTTGTATTAAGTCATCAA	955
QY	860	ctgactcgttgaatcagagcgttccaaggtggaagatcattagctatataataaagccttt	919
Db	956	CTGACTGGTGATTCACAGGCTGTTCACAGGAAGAGCTTCTGAGCTGTATTAAGGCTTTC	1015
QY	920	taccatcttgycttgaagatg	939
Db	1016	TACCTCTCGCTGAGATG	1035

AX061217	AX061217	1726 bp	DNA	PAT	22-JAN-2001
LOCUS	AX061217	Sequence 64 from Patent WO00/89953.			
DEFINITION	AX061217				
ACCESSION	AX061217				
VERSION	AX061217.1	GI:12406353			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 1726)				
	Lai, P., Yang, J., Yue, H., Hillman, J. L., Tang, Y. T., Bandman, O.,				
	Burford, N., Baughn, M. R., Azimzai, Y., Lu, D. A., Au-Young, J. and				
	Patterson, C.				
TITLE	Human transport proteins				
JOURNAL	Patent: WO 0078953-A 64 28-DEC-2000;				
FEATURES	Incyte Genomics, Inc. (US)				
source	1..1726				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	552 a 302 c 399 g 473 t				
ORIGIN					
Query Match	65.1%;	Score 676.8;	DB 9;	Length 1726;	
Best Local Similarity	99.7%;	Fred. No. 1.9e-181;			
Matches 678; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY 332	acgtgctgattcttgagagtcgacgtgcatcatggaacatccgagagtggtgttg	391			
DB 452	acccagtgattcttgagagtcgacgtgcatcatggaacatccgagagtggtgttg	511			
QY 392	gcaaaagtgaaagtgaacatcatcccttggaaatcagltcgtggaaggaatgatgctg	451			
DB 512	gcaaaagtgaaagtgaacatcatcccttggaaatcagltcgtggaaggaatgatgctg	571			
QY 452	ggttatgttgccagtttttaagcaatccaactgacctgtgaagtttcaatgaacaaagg	511			
DB 572	ggttatgttgccagtttttaagcaatccaactgacctgtgaagtttcaatgaacaaagg	631			
QY 512	aaggaaaaaagaacttggaaggaacaaatcgtgatttcgtgtgtacatcatgcatctg	571			
DB 632	aaggaaaaaagaacttggaaggaacaaatcgtgatttcgtgtgtacatcatgcatctg	691			
QY 572	caaaaatcttaagctgaaaggaagaaatcagaggtcttgagcaggtggtaaccaatatac	631			
DB 692	caaaaatcttaagctgaaaggaagaaatcagaggtcttgagcaggtggtaaccaatatac	751			

[p://ftp.genome.washington.edu/RM/RepeatMasker.html](http://ftp.genome.washington.edu/RM/RepeatMasker.html)
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

1	1109	: gap of 1109 bp in length
1110	1209	: gap of 100 bp in length
1210	3242	: contig of 2033 bp in length
3243	3342	: gap of 100 bp in length
3343	4984	: contig of 1642 bp in length
4985	5084	: gap of 100 bp in length
5085	8149	: contig of 3065 bp in length
8150	8249	: gap of 100 bp in length
8250	10485	: contig of 2236 bp in length
10486	10585	: gap of 100 bp in length
10586	12579	: contig of 1994 bp in length
12580	12679	: gap of 100 bp in length
12680	15104	: contig of 2425 bp in length
15105	15204	: gap of 100 bp in length
15205	19613	: contig of 4409 bp in length
19614	19713	: gap of 100 bp in length
19714	23257	: contig of 3584 bp in length
23298	23397	: gap of 100 bp in length
23398	27445	: contig of 4048 bp in length
27446	27545	: gap of 100 bp in length
27546	33569	: contig of 6024 bp in length
33570	33669	: gap of 100 bp in length
33670	39866	: contig of 6217 bp in length
39867	39986	: gap of 100 bp in length
39987	46130	: contig of 6144 bp in length
46131	46230	: gap of 100 bp in length
46231	53403	: contig of 7173 bp in length
53404	53503	: gap of 100 bp in length
53504	61709	: contig of 8206 bp in length
61710	61809	: gap of 100 bp in length
61810	70390	: contig of 8581 bp in length
70391	70490	: gap of 100 bp in length
70491	83671	: contig of 13181 bp in length
83672	83771	: gap of 100 bp in length
83772	99055	: contig of 15264 bp in length
99056	99135	: gap of 100 bp in length
99136	114877	: contig of 15742 bp in length
114878	114977	: gap of 100 bp in length
114978	132866	: contig of 17889 bp in length
132867	133966	: gap of 100 bp in length
132967	157662	: contig of 24696 bp in length
157663	157766	: gap of 100 bp in length
157766	185614	: contig of 27852 bp in length

```
FEATURES      Location/Qualifiers
      source    1..185614
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="18"
```

[illegible]

RESULT	8			
LOCUS	AC008104			
DEFINITION	185614 bp DNA	HTG	26-MAY-2000	
ACCESSION	AC008104			
VERSION	AC008104.3			
KEYWORDS	GI:8072415			
SOURCE	HTG; PHASE1; HTGS-DRAFT.			
ORGANISM	human.			
REFERENCE	1 (bases 1 to 185614)			
AUTHORS	Blumen, B., Linton, L., Nussbaum, C. and Lander, E.			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 185614)			
AUTHORS	Blumen, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckler, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeRellano, K., DePatre, E., Devon, K., Dewar, K., Donegan, L., Doyle, M., Ferrer, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gargay, S., Gilbert, D., Grant, G., Hagos, B., Heath, A., Horton, L., Howland, J., Jones, C., Kann, L., Karas, A., Lebeck, J., Lien, C., Locke, C., MacDonald, P., Margulis, N., McEwan, P., McGuck, M., McKernan, K., McLaughlin, J., Meldrum, J., Molla, M., Morris, M., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Teste, S., Torrella-Miller, L., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-JUL-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
COMMENT	On May 25, 2000 this sequence version replaced g1:7321473. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://f.gene.washington.edu/RM/RepeatMasker.html			
	Genome Center			
	Center: Whitehead Institute/ MIT Center for Genome Research			
	Center code: MIBR			
	Web site: http://www-seq.wi.mit.edu			
	Contact: sequence_submissions@genome.wi.mit.edu			
	Project Information			
	Center project name: L1022			
	Center clone name: 446.F.17			
	Summary Statistics			
	Sequencing vector: M13; W77815; 92% of reads			
	Sequencing vector: Plasmid; n/a; 8.0-14% of reads			
	7.56620428751576chemistry: Dye-terminator-amerham; 92% of reads			
	Chemistry: Dye-terminator Big Dye; 8% of reads			
	Assembly program: Phrap; version 0.960731			
	Consensus quality: 166796 bases at least Q40			
	Consensus quality: 174175 bases at least Q30			
	Consensus quality: 177996 bases at least Q20			
	Insert size: 192000; agarose-fp			
	Insert size: 183514; sum-of-contigs			
	Quality covera.			
	NOTE: This is a 'working draft' sequence. It currently			
	consists of 22 contigs. The true order of the pieces			
	is not known and their order in this sequence record is			
	arbitrary. Gaps between the contigs are represented as			
	* runs of N, but the exact sizes of the gaps are unknown.			
	* This record will be updated with the finished sequence			
	* as soon as it is available and the accession number will			
	* be preserved.			
	I 1109: contig of 1109 bp in length			
	I 1110 1209: gap of 100 bp			
	I 1210 3242: contig of 1003 bp in length			


```

* 3243 3342: gap of 100 bp
* 3343 4984: contig of 1642 bp in length
* 4985 5084: gap of 100 bp
* 5085 8149: contig of 3065 bp in length
* 8150 8249: gap of 100 bp
* 8250 10485: contig of 2236 bp in length
* 10486 10585: gap of 100 bp
* 10586 12579: contig of 1994 bp in length
* 12580 12679: gap of 100 bp
* 12680 15104: contig of 2425 bp in length
* 15105 15204: gap of 100 bp
* 15205 19613: contig of 4409 bp in length
* 19614 19713: gap of 100 bp
* 19714 23297: contig of 3584 bp in length
* 23298 23397: gap of 100 bp
* 23398 27445: contig of 4048 bp in length
* 27446 27545: gap of 100 bp
* 27546 33569: contig of 6024 bp in length
* 33570 33669: gap of 100 bp
* 33670 39886: contig of 6217 bp in length
* 39887 39986: gap of 100 bp
* 39987 46130: contig of 6144 bp in length
* 46131 46230: gap of 100 bp
* 46231 53403: contig of 7173 bp in length
* 53404 53503: gap of 100 bp
* 53504 61709: contig of 8206 bp in length
* 61710 61809: gap of 100 bp
* 61810 70390: contig of 8581 bp in length
* 70391 70490: gap of 100 bp
* 70491 83671: contig of 13181 bp in length
* 83672 83771: gap of 100 bp
* 83772 99035: contig of 15264 bp in length
* 99036 99135: gap of 100 bp
* 99136 114877: contig of 15742 bp in length
* 114878 114977: gap of 100 bp
* 114978 132866: contig of 17889 bp in length
* 132867 132966: gap of 100 bp
* 132967 157662: contig of 24696 bp in length
* 157663 157762: gap of 100 bp
* 157763 185614: contig of 27852 bp in length.

```

FEATURES

```

source
1. 185614.
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone.lib="RP11-446F17"
1. 1109
/note="assembly-fragment"
misc_feature
1210..3242
/note="assembly-fragment"
misc_feature
3343..4984
/note="assembly-fragment"
misc_feature
5085..8149
/note="assembly-fragment"
misc_feature
8250..10485
/note="assembly-fragment"
misc_feature
10586..12579
/note="assembly-fragment"
clone_end:sp6
vector_side:right"
12680..15104
/note="assembly-fragment"
misc_feature
15205..19613
/note="assembly-fragment"
misc_feature
19714..23297
/note="assembly-fragment"
misc_feature
23398..27445
/note="assembly-fragment"
misc_feature
27546..33569
/note="assembly-fragment"
misc_feature
33670..39886

```

```

misc_feature /note="assembly-fragment"
39987..46130
/note="assembly-fragment"
clone_end:r7
vector_side:right"
46231..53403
/note="assembly-fragment"
53504..61709
/note="assembly-fragment"
61810..70390
/note="assembly-fragment"
70491..83671
/note="assembly-fragment"
83772..99035
/note="assembly-fragment"
99136..114877
/note="assembly-fragment"
114978..132866
/note="assembly-fragment"
132967..157662
/note="assembly-fragment"
157763..185614
/note="assembly-fragment"
misc_feature
BASE COUNT 56535 a 35534 c 35072 g 56364 t 2109 others
ORIGIN

```

```

Query Match 12.9%; Score 134; DB 60; Length 185614;
Best Local Similarity 96.5%; Pred. No. 7.1e-27;
Matches 137; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 9 ggtccgttatgctttgctgctactgctgcatgctgcctccggaagagagagagct 68
|||||
Db 50849 GGAGTGCATTATCGTTGGCTACTGCTGAATGCTCCCGAGAGAGAGAGGCT 50908

QY 69 ttggccgtaccacagatagtcgccgagacgaacttcactgtccggtcgccggc 128
|||||
Db 50909 TTGGCCGCTGACCCACAGAGAGGCCCGAGAGAGCAATTCCTACTCTCCGCTGCCGGC 50968

QY 129 taccgtggccgagctagaacc 150
|||||
Db 50969 TACCGTGGCCGAGCTAGTACC 50990

```

```

RESULT 9
AC007576 112126 bp DNA PLN 16-JUN-1999
LOCUS AC007576/c
DEFINITION Arabidopsis thaliana chromosome I BAC F7A19 genomic sequence,
complete sequence.
ACCESSION AC007576
VERSION AC007576.3 GI:5019264
KEYWORDS HTG.
SOURCE
ORGANISM Arabidopsis thaliana
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 112126)
REFERENCE
Fedorispel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altati,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Unpublished
2 (bases 1 to 112126)
REFERENCE
Fedorispel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altati,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (17-MAY-1999) DNA sequencing and technology center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,

```

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 112126)
USA
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (08-JUN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
4 (bases 1 to 112126)
Federpiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (16-JUN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Jun 8, 1999 this sequence version replaced gi:4887737.
e-mail for correspondence: arabes@sequence.stanford.edu
Genes with similarity to proteins in the databases are described as
'putative', '-like' or 'similar to'. Genes that have EST
similarity but no significant protein similarity are described as
'unknown proteins'. Genes that are annotated based only on gene
prediction software are described as 'hypothetical proteins'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
http://genome.stanford.edu/~chris/GENSCAN.html), Fexa (V.Solovay
& A.Salanov, Sanger Centre, http://genomic.sanger.ac.uk/),
NetPlantene (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/NetPlantene.html) and
eMOTif (Newill-Manning, C.G. Wu, T.D. & Brutlag, D.L.,
http://motif.stanford.edu/projects.html).
Location/Qualifiers
1..112126
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/clone="F7A19"
/chromosome="1"
/map="92358"
1240..2806
/gene="F7A19.1"
join(1240..1265,1351..1448,1540..1697,1788..1887,
1976..2354,2426..2507,2612..2806)
/gene="F7A19.1"
/note="Hypothetical protein"
/codon_start=1
/protein_id="A039278.1"
/db_xref="GI:5080768"
/translation="MDLIRKTRKHPASQKIREVSSSTREKKVARKSVSEKEDK
KPSNWLKQKSRQMSQSTYPIGMDAAVAATAYATATFEETLWLNTHSGELGGS
SSRSKSRSEELPLEEPRLSRFSQSLSTIDSTKDHKPTLKKPARKSSVYKITS
MNLMDGTHKONODEEKEHROKRPVSEPPRIOPPLRSEPPRAPPPPLSLSPAR
LPRETRROSEHSTRKSDSTADAMEAKELSKIKARLEKLRKIDLEAKREKARAK
LDISESELEORRRKRGLOFRREDTEYIEQIAAGRAQAERDROSKKEKVKACVITS
TGKLPNACPF"
TGLTGNACPF"
4106..4769
/gene="F7A19.2"
4106..4769
/gene="F7A19.2"
4143..4610
/gene="F7A19.2"
/note="Unknown protein; location of ESTs 128A14T7,
gb|U44625 and 128A14Xp 3', gb|AI099841"
/codon_start=1
/protein_id="A039279.1"
/db_xref="GI:5080769"
/translation="MNFISQVKKLSSSTPEEDPHNKPEVETETATRAPTAELMASA
KVAEAQAARNESEKLDKGVKVASADILDAEAKYCKDEKSTGQYIDKAEKLYIN
DYESSHSTGAGGPPPTSQAPASQPPAPKADDESGGLGAYAKMAQGLK"
complement(5755..8975)
/gene="F7A19.3"
complement(join(5755..7820,7930..8641,8781..8975))
/gene="F7A19.3"
/note="Hypothetical protein"
/codon_start=1
/protein_id="A039280.1"
/db_xref="GI:5080770"
/translation="MGKRVHCESLFGYHHSMGLDKNESNGGRMPLFYGVNDTSASN
NDCCYNGPSTQTFEGFDKDYVLEIHHVYETQKDMDELKRPQNKEMVQTEASTSQ
ATNDVYKAKIPSPPLANSYDRPSMSVVDNGHSPKSGNSGPPVSMQGASKSVY
VSEVYRPIKIRKMDILCPDEYITDNEEVELEKDHVCESTSQLPBGVKTSTRIDG
VRVGGSSRNGADLNEPDAQENAEAGSHSDIENGEFQSHIRLQYKSLNSGVYR
EHIEVILPGLDLNPKPKVMSHOPJRTDYHGTIKRSAPFOPAPKPLDLSQPMOVLN
SSORVMSLPNGSPSPKAVVWRBERFTLDADTONSHAEVITHONHDSLSLHOOP
LYPIYRDSAVSNHLLHSNONSFGPQVYASAKRPVYLMMDTLGNKQGYLGD
RLQPSNRFDSGCGNSRLNHNFTNECTSSKAVAGTGYTPNGGRSDYSERK
VRDLNLTAVTNSVSVYVRKDEHLATLPLMLNPKVCSNELADGKWSLSDAVPS
SLKPLDIRDEAGKQVONIMWLERLKSQCSNNVTEKIDANKETPGIAYDQGNVYR
KVHYRMLDMNEPCDPLMDDOQTEQOTETRVASNNCOIDLNLGSDGKNSVYA
SSRLNSKAPMIDLETVEPESDDEEDGNISEKRSYKSIIEKTEKPEPEKLAETI
VAIYSAQIDREVEYVASSVSETILLMFAEYVNTHEKNDKDLTFRNSQSIDBI
DYRESMTLQLPDISQETMRKPLPEMLKLEETGTALVTSQRRGNARKGQSRDF
QRIIDLPGLSLSHAEVTDIQMDGFEARAGVSTPGLTKRTGSGRRRAITTPA
EPVYCPVAPAPPPSVYQGVHVSQNSNGEMCLEDERSFAGMGMKTRPRRORCPSSS
TTTATSTNQSSSHAPPT"
16683..17720
/gene="F7A19.4"
/note="Eif-5A"
join(16683..16802,17074..17196,17290..17346,17420..17494,
17619..17720)
/gene="F7A19.4"
/note="Eif-5A"
/codon_start=1
/product="initiation factor 5A-4"
/protein_id="A039281.1"
/db_xref="GI:5080771"
/translation="MSDEHHFESSDAGASKTYPOAGTIRKNGYIYIKNRPKVVEY
STSKTGKHGAKCFEVALIDFTSKLSDIYSSHNCVDPVNTVDQILDISDGYVS
LTDNGSTKDDKDLPNDDTLQDKSFDGKRLVYSVMSAMEEQINAKDQGRK"
19674..22095
/gene="F7A19.5"
join(19674..20007,20763..21468,21551..21715,21837..22095)
/gene="F7A19.5"
/note="Similar to DNA-binding proteins"
/codon_start=1
/protein_id="A039282.1"
/db_xref="GI:5080772"
/translation="MSEKEEAPSTSKSTGASRPRLSLPRPSEMFNFNGVGFSPRP
MTLVSNMFPPDSDEFERFSQLAGAMSSPAATAAATAASDQRLDEGTNSSGDVD
PRFKONRPVLDICNIYOGSYGMTHQVLAQVAAVANOANMOPOTYPPPSQVO
SFSSGGOPIPTSAIPAPORERTSDYTIIEHSOOLPANDKADGGMKRYKQGVGS
EPRRSYKCTNPGCPVYKATYRSIDQVYITIKQGNHEPPOPTNRKGNKONTANING
SSINNNGSSSELASQFOTVNSMKTREOHEAVSQATTEHLSADSDGEGVNGETDV
KVEDDNDPDKRSTREYISEPAPASHRTVITVYTTSSVDLIDLDGQYRRKRGQ
KYVGNGYPPSYKCTPGCGVGRHVRATIDKAVVYTEGKHNDLPAKSSSHAA
AAALRDNRNPGGLANLNOOQOOPVAVRLKKEQTT"
complement(22794..24934)
/gene="F7A19.6"
complement(join(22794..22904,23002..23124,23218..23412,
23491..23791,23882..24020,24112..24381,24466..24552,
24646..24934))
/gene="F7A19.6"
/note="Hypothetical protein"
/codon_start=1
/protein_id="A039283.1"
/db_xref="GI:5080773"
/translation="MGGCVSQRKLSNKLQOKKRGSRGSKSKISASMDVPMKRM
SNASVDRFANSEYVHLDPDNCAKMKKRAEMSNANFHLTQLOMNSQIDGNISHEE
AWYDSFYIDSDSDGSSSVFEDANASAGVIOYEFEESTYLKIDGNKGGEYSSKN
EVSIRKQVADSEHHEFTTTTCEHDHRRKSSKVVVMSVRSSTIDSKSTPSEFSSA

Db 620 CTTTATCAAGAAATACCTGCTGGCTTACACAGCTGATTCAGCATTTATAGTGGCAA 679
 QY 476 atcaactgaactagtgaaaggttcagatgcacaaatgaaggaagaaactggaagaa 535
 Db 680 ATCCACATGACCTGTTAAAGTTCGACTCAATCTGAAGTAA---CTCCCCCTGGGG 736
 QY 536 aaccatgcgattctgtgtgtacatcagcttgcgaagaactagctgaagggaa 595
 Db 737 TACCGAGACGTTATTCAGGGGGCGCTAAATGCTTATTCACCATATGCAAAAAGGAAGGAC 796
 QY 596 taagaaggtcttggaaggtctggtacccaatatacaagaagcagcagctgtaataatg 655
 Db 797 TTGCTGCTGTGGAGCTGGGCTTGGCTTAATTTGCCGCAATGCTATTTAATGCTG 856
 QY 656 gagattaccactatgatgaagtgaaacactactggttattgaatacaactgaag 715
 Db 857 CTGAATTTGGCCAGTTATGATCAAGTGAACAGACAACTTTAAATTTACAGAGTCTCAG 916
 QY 716 acaatatacagctcagcttatacaagttatgtcttgagctgagcttcatctcgtg 775
 Db 917 ATATATTTTTCATATTTTACCGGCTGCGGAGCTGTTTTCCTGCTGATCG 976
 QY 776 gaacaccagcagtcacataaagcagataatgaatcaacacagagataaagaagaa 835
 Db 977 GTTCTGCTGTGATGATGAAGTCAAGATGATGGAGATTCAGCTTCAAAAAGCACAT 1036
 QY 836 gggagacttctgataaatacagactgactgcttgaatcaagctgtca 884
 Db 1037 TTGATTTTTCATCAAGACATGAAATGATGGCTTCTGCTTTTGA 1085

RESULT 11
 LOCUS STUCPMRNA 1434 bp mRNA PLN 12-SEP-1997
 DEFINITION Solanum tuberosum mRNA for mitochondrial uncoupling protein.
 ACCESSION Y11220
 VERSION Y11220.1 GI:2398828
 KEYWORDS mitochondrial uncoupling protein; UCP gene.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 Asteridae; Solanales; Solanales; Solanales; Solanaceae; Solanum; Potatoe;
 section Petota.
 1 (bases 1 to 1434)
 Laioi, M., Klein, M., Riesmeier, J.W., Muller-Rober, B., Fleury, C.,
 Bouillaud, F. and Riquier, D.
 A plant cold-induced uncoupling protein
 Nature 389 (6647), 135-136 (1997)
 97441051
 2 (bases 1 to 1434)
 Klein, M.
 Direct Submission
 Submitted (13-FEB-1997) M. Klein, MPI fuer Molekulare
 Pflanzenphysiologie, AG Mueller-Roeber, Karl Liebknecht Str. 25,
 Haus 20, 14476 Golm bei Potsdam, FRG
 Location/Qualifiers
 1. 1434
 /organism="Solanum tuberosum"
 /cultivar="desiree"
 /db_xref="taxon:4113"
 /clone="StuCP7"
 /clone_13b="pot-flower-zap"
 187. 1107
 /codon_start=1
 /evidence="experimental"
 /product="mitochondrial uncoupling protein"
 /protein_id="CA472107.1"
 /db_xref="GI:2398828"
 /db_xref="SPRMBL:024391"
 /translation="MGGDHGKSDISFAGIFASAPAEACTLLPLDRAKYLIO
 OKRAVEDGDLALPKYRGLIGVCTIAKSEFIAKMGIVPILHOCYGLRIGMYEP
 VKNLYGKDHVGDPLSKITLALFTGALGTTIANPDLVAVRLQAEKGLPAGVPRRI

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 Source

CDS

SCALNASTYITKOESEVRLMTGLGNIGRNAINAELASYDQKEAVLRIGPTDNV
 VTHIAGLAGAFPAVICSPVDVKSRRMGDSAXKNTLDCEVTKLKNQPLAFYKGF
 PNFGRIGSWNVIMFLELOAKFKVKSLESP*
 BASE COUNT 393 a 258 c 315 g 468 t
 ORIGIN

Query Match 11.2%; Score 116.2; DB 15; Length 1434;
 Best Local Similarity 49.1%; Pred. No. 5.1e-22;
 Matches 375; Conservative 0; Mismatches 373; Indels 15; Gaps 2;

QY 97 gcgagcaattccctactgtcgcgcgtgacccgtgagccagcttagcaacttccc 156
 Db 229 GCCGGAATATTCGCAAGTATGCTGCTTGTGCTGAGGCGGTACTTACCA 288
 QY 157 ctgagctcacaagaactcagctccacaagaagaagcgcctctgtcgtttgga 216
 Db 289 TTGGATACTGCTAAGTATGACTTCAAGTCAAAAGAGGACG-----TTGAA 336
 QY 217 gacgtgcaagaagatctgccccctataagggaaatgtgcacagccctaggaatc 276
 Db 337 GGGGATGGGCTTAGCTTTACTTAATATAGGATTTATAGTACTGTGGCACATTGCA 396
 QY 277 gaagagaaggtcttctaaagcttggcaagagtgacaccccgcaattacagacgta 336
 Db 397 AAGGAAGAAGAAATAGCTTCACTATGGAAGGATACGTACGTGGTTACATCGTCAATGT 456
 QY 337 ggtatctcggaggtgcgaatgtgtacatctgaacatctccggaaggtgtgttgcgaa 396
 Db 457 ATATATGAGAGCTCTTCGGATTGGGATGGAACCTGTTAAACCTTATGTTGGCAAA 516
 QY 397 agtgaagatgaagatcattccctctggaaatcagtcattgaggaatgaatgtgtgt 456
 Db 517 GATCATGTTGGGATGTGGCATTTGCAAGAAATATCTTGCGCACTTACATCGTGGCG 576
 QY 457 attggccagcttcttgccaatcccaactgactagtgaaaggttcagatgcgaatgaa 516
 Db 577 TTGGGCTATTCATCAATTCCTACATCTACATCTAAGTAAAGTACGTCTTCAAGTGA 636
 QY 517 aaaaagaacttgaaaggaagaacattgagattctgtgtgtacatcagatcttgcaaa 576
 Db 637 AATATG---CAAGCAGGCTGTGCGAGGCTTATTTGAGGCTTAAAGTCCACTCAACA 693
 QY 577 atctagctgaaggaagaaatcagaggtcttggaagctggtggaacccaataaaga 636
 Db 694 ATAGGAAGAAACAGAGAGAGAGTGTGAGCTCTGTGACTGTGAGCCCAATATTGGGCGG 753
 QY 637 gcaagcactggtgataatgagagatttaaccactatgatagatgaagaaactactgtga 696
 Db 754 AATGCCATCATCAATGACAGCTGAATTTACAGTATGATCAAGTGAAGAGAGCTGTTCTT 813
 QY 697 ttgaatacaccactggaagacataatcatgactcagcttatacaagttatgtctgtga 756
 Db 814 AGGATTTCTGGGTTCAACAGCAATGTTGTACTCTTTGATTTGCGGGCTTGGAGCTGT 873
 QY 757 ctggtactctctatctgtggaacccagcagatgcatcacaagaagaagataatgaa 816
 Db 874 TTTTTCGAGTTTGGATAGGCTCTCTGTATGTGTGTAAGTGCAGAGATATGGGAAT 933
 QY 817 ccaagagataaacaagaagaggtcttgtataaataatcga 859
 Db 934 TCCGCATACAAAATATCTTGATGTTTGTTCANAAACATTGA 976

RESULT 12
 LOCUS AC009216 169457 bp DNA HTG 03-DEC-1999
 DEFINITION Drosophila melanogaster chromosome X clone BACR3BH23 (D900) RPII-98
 38.H.23 map 16D-16F strain Y; cn bw sp, *** SEQUENCING IN PROGRESS
 ACCESSION AC009216
 VERSION AC009216.8 GI:6513905

KEYWORDS HTG; HTGS_PHASE1.

REFERENCE 1 (bases 1 to 169457)
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

```

* 71607 76897: contig of 5291 bp in length
* 76898 76977: gap of unknown length
* 76978 83732: contig of 6755 bp in length
* 83733 83812: gap of unknown length
* 83813 89149: contig of 5337 bp in length
* 89150 89229: gap of unknown length
* 89230 97652: contig of 8423 bp in length
* 97653 97732: gap of unknown length
* 97733 106244: contig of 8512 bp in length
* 106245 106325: gap of unknown length
* 106326 116692: contig of 10368 bp in length
* 116693 116772: gap of unknown length
* 116773 129843: contig of 13071 bp in length
* 129844 129923: gap of unknown length
* 129924 130419: contig of 496 bp in length
* 130420 130499: gap of unknown length
* 130500 131145: contig of 646 bp in length
* 131146 131225: gap of unknown length
* 131226 131840: contig of 615 bp in length
* 131841 131920: gap of unknown length
* 131921 132549: contig of 629 bp in length
* 132550 132629: gap of unknown length
* 132630 133129: contig of 500 bp in length
* 133130 133209: gap of unknown length
* 133210 133458: contig of 249 bp in length
* 133459 133538: gap of unknown length
* 133539 134252: contig of 714 bp in length
* 134253 134332: gap of unknown length
* 134333 134855: contig of 523 bp in length
* 134856 134935: gap of unknown length
* 134936 135396: contig of 461 bp in length
* 135397 135476: gap of unknown length
* 135477 135988: contig of 512 bp in length
* 135989 136068: gap of unknown length
* 136069 136855: contig of 787 bp in length
* 136856 136935: gap of unknown length
* 136936 137763: contig of 828 bp in length
* 137764 137843: gap of unknown length
* 137844 138304: contig of 461 bp in length
* 138305 138385: gap of unknown length
* 138386 138905: contig of 521 bp in length
* 138906 138985: gap of unknown length
* 138986 139626: contig of 641 bp in length
* 139627 139706: gap of unknown length
* 139707 140425: contig of 719 bp in length
* 140426 140505: gap of unknown length
* 140506 141245: contig of 740 bp in length
* 141246 141325: gap of unknown length

```

Query Match 11.28: Score 116; DB 61: Length 169457;
 Best Local Similarity 54.48; Pred. No. 9.3e-22;
 Matches 233; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```

OY 415 ccccttggaaatcaatgcatgagggatgagtgctgtgtattgagcaatttttacc 474
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122413 CCGGTTTGGAGTGCGCCCTTGCGGCTGACGGCGGAGCCGTTGCCACTGCTTGC 122472
OY 475 aatcaactgactagtgaaaggttcaagatgaatgaagaaagaaagaaactggaaga 534
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122473 TCGCCGCTGACCTGCTCAAGTGCAATTCGAATGAGGCGCGACGCTGATGCGC 122532
OY 535 aaacattgagattcgtgtgtacatacatgacattgcaaaatcttagctgaagagga 594
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122533 GAGCCACCGAGGCTGACATCCGCTGCGCATCTTCCGCCAGATGTCACGTCGCGGA 122592
OY 595 ataagagagcttggagcaggtgtgttaccatatacaaaagcagcactgtgtaattg 654
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122593 ATTAAAGGCGCTGTGGAAGGAGGATCCGGAATGTCAGCGCGCTTGTGCAATCTG 122652
OY 655 ggaagatttaacacttatgatacagtgaaacactactgtgattgaaatacaacacttgag 714
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122653 GGGACACCTTAACACATACGACATCAAGCACTGATCATGATCGCTGCAGATGCC 122712

```

```

OY 715 gacaaatcaatgactacaggttatacaagttatgttcgtgactgtgattcttacttcg 774
|| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122713 GACTGTACACAGACAGTACGCTGCTGCGCTCTGTTGTGCGGATTCGCGACGATCATCG 122772
OY 775 ggaacaccagcgcgtatgtatataaagcaggaatatgatacaccagagataaacaaga 834
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122773 GGCACGCCAGCTGATGTGTGAAGCGCGCATCATGAACCCACGACGAGCAATGCGC 122832
OY 835 agggagact 842
|| || |||||
Db 122833 CGGTAAGT 122840

RESULT 13
AC012162/c
LOCUS
DEFINITION
Drosophila melanogaster chromosome X clone BACR01N10 (D1115)
PCRT-98 01.N.10 map 16D-16F strain Y: cn bw sp, *** SEQUENCING IN
PROGRESS ***, 142 unordered pieces.
ACCESSION
AC012162
VERSION
HMG: HTGS PHASE1.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 191504)
Celisner,S.E., Agbayan,A., Arcaluna,T.T., Baxter,E., Blazey,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,
Woollay,P., Yang,S., Yee,M., Yu,C. and Rudin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 191504)
Celisner,S.E., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Klm,E., Lee,B., Lewis,S., Li,P., Lomoloni,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Svirskas,R.R.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rudin,G.M.
Direct Submission
Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 11, 2000 this sequence version replaced q1:6838825.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 142 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
457 456: contig of 456 bp in length
* 537 536: gap of unknown length
* 1005 1084: gap of unknown length
* 1085 1084: gap of unknown length
* 1959 2038: gap of unknown length
* 2039 2248: contig of 210 bp in length
* 2249 2328: gap of unknown length

```

2329 2858: contig of 530 bp in length
2859 2938: gap of unknown length
2939 3325: contig of 387 bp in length
3326 3405: gap of unknown length
3406 4165: contig of 760 bp in length
4166 4245: gap of unknown length
4246 4716: contig of 471 bp in length
4717 4796: gap of unknown length
4797 5373: contig of 577 bp in length
5374 5453: gap of unknown length
5454 6050: contig of 597 bp in length
6051 6130: gap of unknown length
6131 6720: contig of 590 bp in length
6721 6800: gap of unknown length
6801 7461: contig of 661 bp in length
7462 7541: gap of unknown length
7542 8104: contig of 563 bp in length
8105 8184: gap of unknown length
8185 8505: contig of 321 bp in length
8506 8585: gap of unknown length
8586 9227: contig of 642 bp in length
9228 9307: gap of unknown length
9308 10091: contig of 784 bp in length
10092 10171: gap of unknown length
10172 11099: contig of 928 bp in length
11100 11179: gap of unknown length
11180 11747: contig of 568 bp in length
11748 11827: gap of unknown length
11828 12767: contig of 940 bp in length
12768 12847: gap of unknown length
12848 13398: contig of 551 bp in length
13399 13478: gap of unknown length
13479 14007: contig of 529 bp in length
14008 14087: gap of unknown length
14088 14867: contig of 780 bp in length
14868 14947: gap of unknown length
14948 15562: contig of 615 bp in length
15563 15642: gap of unknown length
15643 16859: contig of 1217 bp in length
16860 16939: gap of unknown length
16940 17986: contig of 1047 bp in length
17987 18066: gap of unknown length
18067 19341: contig of 1275 bp in length
19342 19421: gap of unknown length
19422 20926: contig of 1505 bp in length
20927 21006: gap of unknown length
21007 22153: contig of 1147 bp in length
22154 22233: gap of unknown length
22234 23236: contig of 1003 bp in length
23237 23316: gap of unknown length
23317 24607: contig of 1291 bp in length
24608 24687: gap of unknown length
24688 25620: contig of 933 bp in length
25621 25700: gap of unknown length
25701 26300: contig of 600 bp in length
26301 26380: gap of unknown length
26381 27488: contig of 1108 bp in length
27489 27568: gap of unknown length
27569 28215: contig of 647 bp in length
28216 28295: gap of unknown length
28296 29440: contig of 1145 bp in length
29441 29520: gap of unknown length
29521 30587: contig of 1067 bp in length
30588 30667: gap of unknown length
30668 32042: contig of 1375 bp in length
32043 32122: gap of unknown length
32123 33740: contig of 1618 bp in length
33741 33820: gap of unknown length
33821 35375: contig of 1555 bp in length
35376 35455: gap of unknown length
35456 36258: contig of 803 bp in length
36259 36338: gap of unknown length
36339 38003: contig of 1665 bp in length

38004 38083: gap of unknown length
38084 39509: contig of 1426 bp in length
39510 39589: gap of unknown length
39590 40719: contig of 1130 bp in length
40720 40799: gap of unknown length
40800 42453: contig of 1654 bp in length
42454 42533: gap of unknown length
42534 43841: contig of 1308 bp in length
43842 43921: gap of unknown length
43922 45410: contig of 1489 bp in length
45411 45490: gap of unknown length
45491 46782: contig of 1232 bp in length
46783 46862: gap of unknown length
46863 48570: contig of 1708 bp in length
48571 48650: gap of unknown length
48651 51042: contig of 2392 bp in length
51043 51122: gap of unknown length
51123 52760: contig of 1638 bp in length
52761 52840: gap of unknown length
52841 54712: contig of 1872 bp in length
54713 54792: gap of unknown length
54793 56521: contig of 1729 bp in length
56522 56601: gap of unknown length
56602 58938: contig of 2337 bp in length
58939 59018: gap of unknown length
59019 60837: contig of 1819 bp in length
60838 60917: gap of unknown length
60918 62968: contig of 2071 bp in length
62969 63068: gap of unknown length
63069 65583: contig of 2515 bp in length
65584 65663: gap of unknown length
65664 68774: contig of 3111 bp in length
68775 68854: gap of unknown length
68855 71235: contig of 2381 bp in length
71236 71315: gap of unknown length
71316 74193: contig of 2798 bp in length
74194 74194: gap of unknown length
74195 77405: contig of 3212 bp in length
77406 77485: gap of unknown length
77486 80936: contig of 3451 bp in length
80937 81016: gap of unknown length
81017 83357: contig of 2341 bp in length
83358 84437: gap of unknown length
84438 87456: contig of 4019 bp in length
87457 87536: gap of unknown length
87537 91532: contig of 3996 bp in length
91533 91612: gap of unknown length
91613 97143: contig of 5531 bp in length
97144 97223: gap of unknown length
97224 105593: contig of 8370 bp in length
105594 105673: gap of unknown length
105674 112071: contig of 6398 bp in length
112072 112151: gap of unknown length
112152 121091: contig of 8860 bp in length
121092 121091: gap of unknown length
121092 143113: contig of 22022 bp in length
143114 143193: gap of unknown length
143194 14387: contig of 394 bp in length
14388 143588: gap of unknown length
143589 143667: gap of unknown length
143668 143927: contig of 260 bp in length
143928 144007: gap of unknown length
144008 144842: contig of 835 bp in length
144843 144922: gap of unknown length
144923 145360: contig of 438 bp in length
145361 145440: gap of unknown length
145441 145915: contig of 475 bp in length
145916 145995: gap of unknown length
145996 146532: contig of 537 bp in length
146533 146612: gap of unknown length

Query Match 11.2%; Score 116; DB 63; Length 191504;
Best Local Similarity 54.4%; Pred. No. 9.4e-22;
Matches 233; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

[illegible]

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	REFERENCE	AUTHORS
AE003506	AE003506 300994 bp DNA of <i>Drosophila melanogaster</i> genomic scaffold 142000013386053 section 23 of 30, complete sequence.	AE003506	AE002593	AE003506.1 GI:7293355	HTG.	fruit fly.	
						<i>Drosophila melanogaster</i>	
						Euarystia, Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota, Endopterygota; Diptera; Brachycera;	
						Muscomorpha; Ephydroidea; Drosophilidae; <i>Drosophila</i> .	
						1 (bases 1 to 300994)	
						Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Mortman,J.R., Vandeil,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor	
						Milos,G.L., Abril,J.F., Agbayani,A., An,H.J.,	
						Xiang,L., Andrews-Plankkoc,H.C., Baldwin,D., Ballew,R.M., Beeson,K.Y.,	
						Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,	
						Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borikova,D.,	
						Borchan,M.R., Bouck,J., Brokstein,P., Brothier,P., Butts,K.C.,	
						Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,	
						Cherry,J.M., Cawley,S., Dalke,C., Davenport,L.B., Davies,P., de	
						Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,	
						Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,	
						Dunn,P., Durkin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S.,	
						Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S.,	
						Gelbart,W.M., Glasser,K., Glodok,A., Gong,F., Gorrell,J.H., Gu,Z.,	
						Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J.,	
						Hernandez,J., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,	
						Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,	
						Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,	
						Kravitz,S., Kulp,D., Lai,Z., Lasco,P., Lei,Y., Levitsky,A.A.,	
						Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Matel,B., McIntosh,T.C.,	
						McLeod,M.P., McPherson,D., Mekulov,G., Milshina,N.V., Mobarry,C.,	

FEATURES	source
TITLE	2 (bases 1 to 300994)
JOURNAL	Adams,M.D., Celisiker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
REFERENCE	Direct Submission
AUTHORS	Submitted (12-MAR-2000) Celera Genomics, 45 West Gude Drive,
JOURNAL	Rockville, MD, USA
source	location/Qualifiers
	1..300994
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/chromosome="X"
mrna	complement(join<16198..16317,16386..16574,16635..16845
	16915..17053,17129..17411,17526..17607,17676..17975,
	18107..>18483>)
	/gene="CG8527"
	/product="CT24901"
	/db_xref="FLYBASE:FBan0008527"
	/db_xref="FLYBASE:FBgn0030844"
	complement(<16198..>18483)
	/gene="CG8527"
	/map="1684..1686"
	/db_xref="FLYBASE:FBan0008527"
	/db_xref="FLYBASE:FBgn0030844"
	complement(join(16198..16317,16386..16574,16635..16845,
	16915..17053,17129..17411,17526..17607,17676..17975,
	18107..18483>)
	/gene="CG8527"
	/note="CG8527 gene product"
	/codon_start=1
	/db_xref="FLYBASE:FBan0008527"
	/db_xref="FLYBASE:FBgn0030844"
	/protein_id="AAF48734.1"
	/db_xref="GI:7293356"
	/translation="MPQERHPPESTRQRFLETLVIRSRSLIYOTKEFFONSLHGR
	VIASGRIGEGEKEMWRCPTSGATVAIVITMSLMEKQOTPTTIGLTDPHNKNVPP
	TYVCPAPADHDHKTVEKYVNTLANTDEAQAQMTPLRLITSLNFENVADAKLLOS
	IPQULDAHTIREMAFEGHIDCKNVFSCATRDIDIPCDHFEPLTYEHGCFANSR
	EKSPTRDKNTKGADHDLYETDKWALFEINRSIRLIFSNEEYFGSDPNAQIDWSPR
	QLVAVTRISKNTYTTDDARQSLSGORCIFSDEKALNVPDAYFFSCMQCRNNKA1
	KLCGNCPEYKLPINVPWCSIKDCEIDFENSTNTIKDCLQCSKVFYNIDKLI
	KMSRPSGLGLVLEFLWPIIRYKREVLDFGVDLVFEGGLASLFLSFLSGVEIYI
	YETIRACCMYKKNQDELXETIEKTRQEPKIDIKLSKHNPRIGDTPISAVLYKPR
	AEEBACPNPLAGIORYKHKEDYSLNTKTYTKTPKVPDYGTTSKTKNSWTTYDHSQYMP
mrna	complement(join(12173..21792,21868..21999,22058..22191,
	22257..22823,22889..23887,24168..24320,24398..25116,
	25191..25449>)
	/gene="CG8465"
	/product="CT24743"
	/db_xref="FLYBASE:FBan0008465"
	/db_xref="FLYBASE:FBgn0030845"
	complement(join(12173..21792,21868..21999,22058..22191,
	22257..22823,22889..23887,24168..24320,24398..25116,
	25191..25449>)
	/gene="CG8465"
	/product="CT24757"
	/db_xref="FLYBASE:FBan0008465"
	/db_xref="FLYBASE:FBgn0030845"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2001, 17:32:31 ; Search time 112.35 seconds
(without alignments)
5398.743 Million cell updates/sec

Title: US-09-397-342-2

Perfect score: 1039
Sequence: 1 ccgagctccgagctccgcttat.....cagatatccatccactgc 1039

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 678276 seqs, 291830651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_0401.*
1: /SID56/gcgdata/geneSeq/geneSeq/NA1980.DAT.*
2: /SID56/gcgdata/geneSeq/geneSeq/NA1981.DAT.*
3: /SID56/gcgdata/geneSeq/geneSeq/NA1982.DAT.*
4: /SID56/gcgdata/geneSeq/geneSeq/NA1983.DAT.*
5: /SID56/gcgdata/geneSeq/geneSeq/NA1984.DAT.*
6: /SID56/gcgdata/geneSeq/geneSeq/NA1985.DAT.*
7: /SID56/gcgdata/geneSeq/geneSeq/NA1986.DAT.*
8: /SID56/gcgdata/geneSeq/geneSeq/NA1987.DAT.*
9: /SID56/gcgdata/geneSeq/geneSeq/NA1988.DAT.*
10: /SID56/gcgdata/geneSeq/geneSeq/NA1989.DAT.*
11: /SID56/gcgdata/geneSeq/geneSeq/NA1990.DAT.*
12: /SID56/gcgdata/geneSeq/geneSeq/NA1991.DAT.*
13: /SID56/gcgdata/geneSeq/geneSeq/NA1992.DAT.*
14: /SID56/gcgdata/geneSeq/geneSeq/NA1993.DAT.*
15: /SID56/gcgdata/geneSeq/geneSeq/NA1994.DAT.*
16: /SID56/gcgdata/geneSeq/geneSeq/NA1995.DAT.*
17: /SID56/gcgdata/geneSeq/geneSeq/NA1996.DAT.*
18: /SID56/gcgdata/geneSeq/geneSeq/NA1997.DAT.*
19: /SID56/gcgdata/geneSeq/geneSeq/NA1998.DAT.*
20: /SID56/gcgdata/geneSeq/geneSeq/NA1999.DAT.*
21: /SID56/gcgdata/geneSeq/geneSeq/NA2000.DAT.*
22: /SID56/gcgdata/geneSeq/geneSeq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1039	100.0	1039	21	AI4084
2	998	96.1	998	21	A37139
3	990.6	95.3	1248	21	Human uncoupling p
4	676.8	65.1	1726	22	F27721
5	286.8	27.6	828	21	C76281
6	140	13.5	1069	21	C46432
7	140	13.5	1071	21	C39993
8	124	11.9	1240	21	C45636
9	124	11.9	1242	21	C40226
10	113.4	10.9	1103	21	C46816
11	106	10.2	1227	21	C33096

12	104.6	10.1	1224	21	C49402	Arabidopsis thalia
13	97.2	9.4	957	21	C49389	Arabidopsis thalia
14	93.4	9.0	876	21	A37719	Novel human protei
15	89.6	8.6	1031	21	A27959	Murine uncoupling
16	88	8.5	1022	21	A27998	Murine uncoupling
17	79.6	7.7	936	21	A28945	Human uncoupling p
18	78.4	7.5	993	21	A27992	Human uncoupling p
19	78.4	7.5	1009	21	A27996	Human uncoupling p
20	78.4	7.5	1674	21	A29803	Human secreted pro
21	77.8	7.5	1178	21	C49924	Arabidopsis thalia
22	77.8	7.5	1180	21	C42453	Arabidopsis thalia
23	76.8	7.3	978	21	A27995	Human uncoupling p
24	76.2	7.3	1416	21	C38589	Arabidopsis thalia
25	74.2	7.1	1411	21	C50232	Arabidopsis thalia
26	63	6.1	1166	21	C34840	Arabidopsis thalia
27	60.8	5.9	1398	20	X25083	Potato tuber-speci
28	57.2	5.5	924	21	A29243	Human uncoupling p
29	57.2	5.5	924	21	A29243	Human uncoupling p
30	56.2	5.4	882	21	A37720	Novel human protei
31	52.6	5.1	1103	21	A27997	Human uncoupling p
32	52.6	5.1	1430	21	A26374	Human secreted pro
33	52.4	5.0	1204	19	V71712	Mouse uncoupling p
34	51.8	5.0	960	22	C83793	Hybrid hUCP2 DNA.
35	50.8	4.9	1658	20	V84307	Mouse uncoupling p
36	50.8	4.9	1949	20	Z07061	Mouse uncoupling p
37	50.8	4.9	1949	20	V71228	CDNA encoding a mu
38	50.8	4.9	1949	21	Z90319	CDNA encoding murti
39	49.4	4.8	930	20	Z19968	Human uncoupling p
40	49.4	4.8	930	20	X99435	UCP2 Nucleotide se
41	49.4	4.8	930	21	Z50624	Tularik human unco
42	49.4	4.8	1105	20	X99434	UCP2 Nucleotide se
43	49.4	4.8	1105	21	Z29323	Human uncoupling p
44	49.4	4.8	1255	17	T13981	Human body weight
45	49.4	4.8	1255	19	V09078	Human C5 gene CDNA

ALIGNMENTS

RESULT	1
AI4084	AI4084 standard; CDNA: 1039 BP.
XX	
AC	AI4084;
XX	
DT	21-JUL-2000 (first entry)
XX	
DE	CDNA encoding human uncoupling protein UCP4.
XX	
KW	Uncoupling protein; UCP4; expressed sequence tag; EST; human;
KW	mitochondrial membrane; ATP synthesis; energy efficiency;
KW	drug screening; obesity; stroke; trauma; sepsis; infection; ss.
XX	
OS	Homo sapiens.
XX	
PH	key
FT	CDS
FT	Location/Qualifiers
XX	/*tag= a
XX	/product= "Human UCP4"
PN	MO200017353-A1.
XX	
PD	30-MAR-2000.
XX	
PF	15-SEP-1999; 99WO-US21194.
XX	
PR	22-SEP-1998; 98US-0101279.
PR	30-DEC-1998; 98US-0114223.
PR	16-APR-1999; 99US-0129674.
XX	
PA	(GETH) GENENTECH INC.
XX	


```

PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.

```

```

PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105593.
PR 26-OCT-1998; 98US-0105594.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105811.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0108500.
PR 03-NOV-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.

```

(GETH) GENENTECH INC.

Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WT;
WPI: 2000-237871/20.
P-PSDB: Y99457.

New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides; useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions

Claim 2; Fig 235; 773pp; English.

A37022 to A37144 encode the new isolated human transmembrane, receptor or secreted PRO polypeptides given in Y99340 to Y99462. The transmembrane and receptor PRO proteins can be used for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoding them have various industrial applications, including uses as pharmaceutical and diagnostic agents. A37145 to A37330 represent PCR primers and hybridization probes used in the isolation of the PRO polypeptides from the present invention.

Sequence 998 BP; 284 A; 202 C; 262 G; 250 T; 0 other;

Query Match 96.1%; Score 998; DB 21; Length 998;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 998; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 14 ccgtatcgtcttgcgtactcgtcgaatgcccgtcccgagagagagagaggtttgc 73
 DB 1 cgttatcgtcttgcgtactcgtcgaatgcccgtcccgagagagagagaggtttgc 60
 QY 74 cgttaaccagagatgccccgagagagagagagagagagagagagagagagagag 133
 DB 61 cgttaccag 120
 QY 134 tggcagagcag 193
 DB 121 tggcagagcag 180
 QY 194 aagcagcgtctcgtcgttggagagagagagagagagagagagagagagagagag 253
 DB 181 aagcagcgtctcgtcgttggagagagagagagagagagagagagagagagagag 240
 QY 254 tggcagagcag 313
 DB 241 tggcagagcag 300
 QY 314 caccgcagcag 373
 DB 301 caccgcagcag 360
 QY 374 tccgag 433
 DB 361 tccgag 420
 QY 434 tggag 493
 DB 421 tggag 480
 QY 494 aggttcag 553
 DB 481 aggttcag 540
 QY 554 ggttaccag 613
 DB 541 ggttaccag 600
 QY 614 ggttaccag 673
 DB 601 ggttaccag 660
 QY 674 ataccag 733
 DB 661 ataccag 720
 QY 734 ggttaccag 793
 DB 721 ggttaccag 780
 QY 794 tcaaaagcag 853
 DB 781 tcaaaagcag 840
 QY 854 cagcag 913
 DB 841 cagcag 900
 QY 914 ggttaccag 973
 DB 901 ggttaccag 960
 QY 974 aaaaaacag 1011
 DB 961 aaaaaacag 998
 RESULT 3
 ID A14085
 XX A14085 standard; DNA: 1248 BP.

AC A14085;
 XX 21-JUL-2000 (first entry)
 DE Human uncoupling protein UCP4 "from DNA" sequence, SEQ ID NO:5.
 XX
 KW Human uncoupling protein; UCP4; expressed sequence tag; EST; UCP3 homologue;
 KW human; chromosome 6p11.2-q12; ATP synthesis; energy efficiency;
 KW mitochondrial membrane; proton leakage; heat production; metabolic rate;
 KW drug screening; obesity; stroke; trauma; sepsis; infection; ds.
 OS Homo sapiens.
 PN WO200017353-A1.
 PD 30-MAR-2000.
 XX 15-SEP-1999; 99WO-US21194.
 PF 22-SEP-1998; 98US-0101279.
 PR 30-DEC-1998; 98US-0114223.
 PR 16-APR-1999; 99US-0129674.
 XX
 PA (GERTH) GENENTECH INC.
 XX Adams S, Pan J, Zhong A;
 PI WPI: 2000-292842/25.
 DR
 XX
 PT New nucleic acid encoding human uncoupled protein-4, useful e.g for
 PT identifying metabolic regulators for treatment of obesity -
 XX
 PS Example 1; Fig 7; 80bp; English.
 CC The invention relates to human uncoupling protein UCP4 (Y81497) and
 CC cDNA encoding it (A14085). The human UCP4 cDNA (ATCC 203134) was isolated
 CC from a brain cDNA library using a probe generated using primers
 CC (A14086-A14087). These primers were based on a UCP4 "from DNA" sequence
 CC (A14085) derived from a number of ESTs (expressed sequence tags) which
 CC were selected on the basis of homology with human UCP3. The human UCP4
 CC gene has been mapped to chromosome 6p11.2-q12. UCP4 catalyses the leakage
 CC of protons through the mitochondrial membrane, thus bypassing ATP
 CC synthase and thereby reducing the efficiency of ATP synthesis. Modulation
 CC of UCP4 activity or expression can therefore alter the metabolic rate and
 CC heat production via modulation of ATP synthetic efficiency. UCP4 nucleic
 CC acids may be used for recombinant production of UCP4 and as a source of
 CC primers and hybridisation probes which may be used for the analysis of
 CC UCP4 expression, for screening for homologous sequences, and for
 CC chromosome or gene mapping. They can be also be used to produce
 CC transgenic or knockout animals for the development and screening of
 CC therapeutic agents, as a source of antisense nucleotides, and in gene
 CC therapy for metabolic disorders. The UCP4 protein can be used to screen
 CC for specific modulators of activity and to raise antibodies which may be
 CC used therapeutically as UCP4 antagonists or UCP4 targeting proteins, for
 CC affinity purification of UCP4, and as immunoassay reagents for detecting
 CC UCP4 expression, e.g., for the diagnosis of impaired neural activity or
 CC neural degeneration. Agents that modulate UCP4 activity are used to
 CC control the metabolic rate in mammals. UCP4 upregulators can be used to
 CC increase the metabolic rate in order to treat obesity and the symptoms
 CC associated with stroke, trauma, sepsis and infection. The present
 CC sequence represents the human UCP4 "from DNA" sequence assembled from EST
 CC sequences with homology to human UCP3.
 XX
 SQ Sequence 1248 BP; 357 A; 263 C; 301 G; 326 T; 1 other:
 Query Match 95.3%; Score 990.6; DB 21; Length 1248;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 993; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 15 cgttatcgtcttgcgtactcgtcgaatgcccgtcccgagagagagagaggtttgc 74
 DB 1 cgttatcgtcttgcgtactcgtcgaatgcccgtcccgagagagagagaggtttgc 60

QY 75 gctaccagagatggtcccgagcagcaaatctcctactgtccgctgctgctacagt 134
 DB 61 gctaccagagatggtcccgagcagcaaatctcctactgtccgctgctgctacagt 120
 QY 135 ggcagcagtagcaaccttccctgagatctcaaaaaactcgactccaatgcaagaga 194
 DB 121 ggcagcagtagcaaccttccctgagatctcaaaaaactcgactccaatgcaagaga 180
 QY 195 agcagccttgctgctggttgagagacgtgtcagaagaatctgccccctataagggaatgt 254
 DB 181 agcagccttgctgctggttgagagacgtgtcagaagaatctgccccctataagggaatgt 240
 QY 255 ggcagcagccttgagatgagatgagagagaggtcttctaagctttggcagaaggtgac 314
 DB 241 ggcagcagccttgagatgagatgagagagaggtcttctaagctttggcagaaggtgac 300
 QY 315 acccgcaattacagacagtagtattctctgaggtcgaatggttcacatagacatct 374
 DB 301 acccgcaattacagacagtagtattctctgaggtcgaatggttcacatagacatct 360
 QY 375 ccgagaggtgtgtttggcaaaagtgaagatgagcatatccctttggaataatgcat 434
 DB 361 ccgagaggtgtgtttggcaaaagtgaagatgagcatatccctttggaataatgcat 420
 QY 435 ttgagagatgagtgctgtgttatttgcccaatttttagcccaactgacagtgaa 494
 DB 421 ttgagagatgagtgctgtgttatttgcccaatttttagcccaactgacagtgaa 480
 QY 495 ggtccagatgcaaatggaagaaagaaactggaagaaacatgcatctgtcgtg 554
 DB 481 ggtccagatgcaaatggaagaaagaaactggaagaaacatgcatctgtcgtg 540
 QY 555 ttgacatcatgcatcttgcacaaatcttaagctgaagaggaatgacaggtgtggcag 614
 DB 541 ttgacatcatgcatcttgcacaaatcttaagctgaagaggaatgacaggtgtggcag 600
 QY 615 ctgggtaccccaatatacaaaagacagcctgtgaatatagtgaatggaatttaaccattatga 674
 DB 601 ctgggtaccccaatatacaaaagacagcctgtgaatatagtgaatggaatttaaccattatga 660
 QY 675 tacagtgaaacactactgtgtatgaaatcacccacttgagagcaaatatcatgactcacg 734
 DB 661 tacagtgaaacactactgtgtatgaaatcacccacttgagagcaaatatcatgactcacg 720
 QY 735 ttatcaagttatgtcttgagactgtgacttctatctctgggaacaccagcgatgcat 794
 DB 721 ttatcaagttatgtcttgagactgtgacttctatctctgggaacaccagcgatgcat 780
 QY 795 caaaagcagaaatgaaatgaaacacagagataaaacaaaggaaggagactttgtataatc 854
 DB 781 caaaagcagaaatgaaatgaaacacagagataaaacaaaggaaggagactttgtataatc 840
 QY 855 atcagactgactgtgactgacagctgttcaaggtgaagatgcatgcatataaag 914
 DB 841 atcagactgactgtgactgacagctgttcaaggtgaagatgcatgcatataaag 900
 QY 915 ctcttaccatctgtgctgagaaatgaaacctgtgtcaatggttctggttactatga 974
 DB 901 ctcttaccatctgtgctgagaaatgaaacctgtgtcaatggttctggttactatga 960
 QY 975 aaatcagagagatgagtgagtcagtcattttaa 1011
 DB 961 aaatcagagagatgagtgagtcagtcattttaa 997

RESULT 4
 F27721
 ID F27721 standard; cDNA; 1726 BP.
 AC F27721;
 XX
 DT 28-MAR-2001 (first entry)

XX XX Human transport protein TPPT-21 coding sequence.
 DE Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer; ss.
 OS Homo sapiens.
 PN W0200078953-A2.
 PD 28-DEC-2000.
 PF 16-JUN-2000; 2000MO-US16668.
 PR 17-JUN-1999; 99US-0139923.
 PR 10-AUG-1999; 99US-0148177.
 PR 18-AUG-1999; 99US-0149357.
 PR 28-OCT-1999; 99US-0162287.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
 PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;
 XX WPI: 2001-041424/05.
 DR P-PSDB: B60101.
 XX Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -
 XX
 PS Claim 5; Page 153-154; 165pp; English.
 CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 CC
 XX Sequence 1726 BP; 552 A; 302 C; 399 G; 473 T; 0 other;
 SQ
 QY 332 acgtatgtatctcgagagtcgaatgtcacatatagcaatctccgagaggtgtgtt 391
 DB 452 acccagtgatctcgagagtcgaatgtcacatatagcaatctccgagaggtgtgtt 511
 QY 392 gcaaaagtgaagatgagcatatcccttggaaatcagtcattgtgagagatgagctg 451
 DB 512 gcaaaagtgaagatgagcatatcccttggaaatcagtcattgtgagagatgagctg 571
 QY 452 gtgtatgtgccaatttttaagccaatccaactgagtcagtgaaagttcaagatgcaatg 511
 DB 572 gtgtatgtgccaatttttaagccaatccaactgagtcagtgaaagttcaagatgcaatg 631
 QY 512 aagaaaaaagaaactggaagaaacacatgcatctgtgtgtacatcatgattg 571
 DB 632 aagaaaaaagaaactggaagaaacacatgcatctgtgtgtacatcatgattg 691
 QY 572 caaaaatctagctgaaggaagaaatgaagggtcttgagcaggtgtggttaccataatc 631
 DB 692 caaaaatctagctgaaggaagaaatgaagggtcttgagcaggtgtggttaccataatc 751
 QY 632 aaagagcagcactgtgtaataatggagatttaaccactatgataaggtgaacactact 691
 DB 752 aaagagcagcactgtgtaataatggagatttaaccactatgataaggtgaacactact 811
 QY 692 ttgtatgaatacaccaacttgagacaatatcatgactcaaggttataagttatgtt 751

DB 812 tggatgtgaataacaccattgagacatacatgactacaggttatcatatgtt 871
QY 752 ctggactggtaagcttctatctctggacacccagctgctcatcaaaagcagataatga 811
DB 872 ctggactggtaagcttctatctctggacacccagctgctcatcaaaagcagataatga 931
QY 812 atcaacacagagataaacaaggaaggagcttctgtataatcatcagactgcttga 871
DB 932 atcaacacagagataaacaaggaaggagcttctgtataatcatcagactgcttga 991
QY 872 tcaagctctcaaggctgaagattcatgactatataaaggctttttaccctctggc 931
DB 992 ttcagagcttctcaaggctgaagattcatgactatataaaggctttttaccctctggc 1051
QY 932 tgaagaatgaccccttggatcaatggtgtctgtgcttactatgaaaaaatacagagatga 991
DB 1052 tgaagaatgaccccttggatcaatggtgtctgtgcttactatgaaaaaatacagagatga 1111
QY 992 gtggagtcagtcattttaa 1011
1112 gtggagtcagtcattttaa 1131

RESULT 5
C76281
ID C76281 standard; cDNA: 828 BP.
AC C76281;
XX
XX
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF1836 polynucleotide sequence SEQ ID NO:3671.
XX
XX
KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulerary; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; diabetes mellitus; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX
Homo sapiens.
PN WO200058473-A2.
XX
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000MO-US08621.
XX
PR 31-MAR-1999; 990US-0127607.
PR 02-APR-1999; 990US-0127636.
PR 05-APR-1999; 990US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI: 2000-602362/57.
XX
XX P-PSDB; BA2072.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
PS Claim 5; Page 2828; 5507pp; English.

XX
CC C74446 to C77606 encode the proteins given in B40237 to B43397, which
CC represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulerary;
CC antiporiatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antirheumatic; immunosuppressant; immunostimulant;
CC dermant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC antiviral; antifungal; antineumatic; antithyroid; antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 828 BP; 276 A; 133 C; 201 G; 216 T; 2 other:

Query Match 27.6%; Score 286.8; DB 21; Length 828;
Best Local Similarity 99.3%; Pred. No. 8.1e-84;
Matches 288; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 332 acgtggtgctatcttgaggtggaatggtgacatggaatccgagaggtgtgtgt 391
DB 539 acccaagtcattctggaggtcgaaatggtgacatggaatccgagaggtgtgtgt 598
QY 392 gcaaaagtgaagatgagcattacccttggaaatcagtcattgagagatgagctg 451
DB 599 gcaaaagtgaagatgagcattacccttggaaatcagtcattgagagatgagctg 658
QY 452 ggttatgtgcaagtttttagccaatcgaactgagctagtgaggttcagatgcaatg 511
DB 659 ggttatgtgcaagtttttagccaatcgaactgagctagtgaggttcagatgcaatg 718
QY 512 aaggaagaaaggaactggaaggaagaaacccttgcatttctgtgttacctgattg 571
DB 719 aaggaagaaaggaactggaaggaagaaacccttgcatttctgtgttacctgattg 778
QY 572 caaaatcttaagctgaaggaagaaatcagaggtcttggcagcggtgta 621
DB 779 caaaatcttaagctgaaggaagaaatcagaggtcttggcagcggtgta 828

RESULT 6
C46432
ID C46432 standard; DNA: 1069 BP.
XX
XX C46432;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 50118.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990US-0121825.

PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125768.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140981.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142920.
 PR 13-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 13.5%; Score 140; DB 21; Length 1069;

Best Local Similarity 51.8%; Pred. No. 1.5e-35; Matches 382; Conservative 0; Mismatches 335; Indels 21; Gaps 2;

QY 279 agaggaagcgttcttaaacgttctgcaagagtgacacccgcatcttacagacgtagt 338
 DB 224 aaaggaagagtgatgtgtctctacaaagtcctacgtcgtcgtatctacagacgtgt 283
 QY 339 gtaattcgagtgcaatggtcacaataatccagag-----aggttggtttgg 392
 DB 284 ctcaagcctatcagatcatttgatcagagatttgaaagcctatcgtcagatctga 343
 QY 393 caaagtgagatgacatcattcccttggaaatcagtcattggaggtatgagctgg 452
 DB 344 aactacaataagcagctcctctcctccgcacaaagcctcgtcgtgaggtttctgg 403
 QY 453 tttatgacagatttttagccaatcgaactgacagtgaaagttcagatgcaatgga 512
 DB 404 tttatagctcagtgatgtgtcgtacgtcagctgatttgcacaaagtgaatgaaagcaga 463
 QY 513 aggaagaaagaaacttgaaagaaaccattgcgattcgtggtatcattcatttgc 572
 DB 464 tggtagattggtgagcgaagcctgaaaccgagtlactcgggccaatcgagcttttac 523
 QY 573 aaaaacttagctgaaaggaatcagaggtcttgagcaggtgtgtatcccaataaca 632
 DB 524 caaaatcccaacaatcgaagaggttaaaggtatgaaaggtgtcttcttccaacaaca 583
 QY 633 aagaagcagcagtgtaataatgaggaatttaacacttatgatacagtgtaaacacact 692
 DB 584 gagagagattctagtgataatgaggaactagctgtctatgatacagccaacaacttgc 643
 QY 693 ggtattgaatcaccaacttgaggaacaataatcagtaactcagcgtgttataagttatgttc 752
 DB 644 catcgataagaagattgctgagagatacaattttgagcacaactcttcttataatgtc 703
 QY 753 tggactggtagcttctatctcggaaacacgagcgtatgcatcaaaagcagataatgaa 812
 DB 704 gggctctgcttcgacaagatttgagttgtccagctgattggtgaaagcagagatgataa 763

QY 813 tcaccacagataaacaagaaggagactttgtataatcactgactgacttgat 872
 DB 764 ccaggtgtgaaatgc-----tgytacagaattcttaagactgtttgt 808
 QY 873 tcaggtcttcaagtgatgaagatcattgagctatataaagcttttccactgtgct 932
 DB 809 gaagaaggttaagtgtgaaagaaagacttgttgaagaggttcttccgacatgggc 868
 QY 933 gagaatgacccttggtcaatggtgtcgtgacttataaanaatcagagagatgag 992
 DB 869 aaggtctggaacccgtgacagcttgcgttttgggtcctatgaaagtttaagactgtgc 928
 QY 993 tggagtcagttcaattta 1010
 DB 929 aggaatccttcttcta 946

RESULT 7
 C39993
 ID C39993 standard; DNA; 1071 BP.
 XX
 AC C39993;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26654.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW Protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 14-MAY-1999; 99US-0134768.
 PR 18-MAY-1999; 99US-0134941.
 PR 19-MAY-1999; 99US-0135124.
 PR 20-MAY-1999; 99US-0135353.
 PR 21-MAY-1999; 99US-0135629.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0135621.

PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140891.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142970.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144864.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 26-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158359.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.

KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX

PK 23-00N-1999; 9905-0140999L.

PR 30-JUN-1999; 99US-014187.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0146708.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0153486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156556.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157173.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 11.9%; Score 124; DB 21; Length 1240;
 Best Local Similarity 53.5%; Pred. No. 2.9e-30;
 Matches 293; Conservative 0; Mismatches 240; Indels 15; Gaps 1;

QY 463 cagttttagccaatccaactgacactgaggttcagatgcgaatggaagaaag 522
 DB 462 caggtatgtgcttagtccaagctgattggtcaagtggaatgcgaagatggtgattg 521
 QY 523 aaactggaagaaacacttgcatlctgltgtgtacatcatgcatlgtgnaaaactla 582
 DB 522 gtgagccaagcgccgaacgcgaggtactgcgacaaatcgaggttttaccnaaatccta 581
 QY 583 gctgaagaggaatacgaaggtcttggaagcgtggttaccnaatacagaagca 642
 DB 582 caatcagaagagtaaaaggtltalgaaaggtlcttccaacacccagagacatt 641
 QY 643 ctgtggaatggaagatttaacacactatgatacagtggaacactctgtgtatgaa 702
 DB 642 ctatgataatggaagactagcttgcataatgcacgccaacacacttgcatacga 701
 QY 703 acacacttgagacaataatcatgatacgaagttatcaagttatgtlctggaactgta 762
 DB 702 aagattgtgagataacattttgcgcacaccttgcctctataatgtcggtgtgtct 761
 QY 763 gcttctattctggaacacacgcgcgaatgtatcaaaagcagaataatgaatcaaccca 822
 DB 762 tcgacaagatttgatgtlccagctgagtggtgtaagacgagatgataacccaaggtgaa 821
 QY 823 gataacaagaagggactttgtataaatcatcgcactgacgtgattgattcaggtgtt 882

[illegible]

PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148365.
PR	16-AUG-1999;	99US-0148684.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151330.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154039.
PR	16-SEP-1999;	99US-0154479.
PR	20-SEP-1999;	99US-0154739.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158332.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159320.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	29-OCT-1999;	99US-0162142.

Query Match

11.98; Score 124; DB 21; Length 1242

[illegible]

PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 16-SEP-1999; 99US-0154779.
 PR 20-SEP-1999; 99US-0155139.
 PR 22-SEP-1999; 99US-0155486.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999: 99US-0159294.
 PR 13-OCT-1999: 99US-0159295.
 PR 14-OCT-1999: 99US-0159329.
 PR 14-OCT-1999: 99US-0159330.
 PR 14-OCT-1999: 99US-0159331.
 PR 14-OCT-1999: 99US-0159637.
 PR 14-OCT-1999: 99US-0159638.
 PR 18-OCT-1999: 99US-0159584.
 PR 21-OCT-1999: 99US-0160741.
 PR 21-OCT-1999: 99US-0160767.
 PR 21-OCT-1999: 99US-0160768.
 PR 21-OCT-1999: 99US-0160770.
 PR 21-OCT-1999: 99US-0160814.
 PR 21-OCT-1999: 99US-0160815.
 PR 22-OCT-1999: 99US-0160880.
 PR 22-OCT-1999: 99US-0160981.
 PR 22-OCT-1999: 99US-0160989.
 PR 25-OCT-1999: 99US-0161404.
 PR 25-OCT-1999: 99US-0161405.
 PR 25-OCT-1999: 99US-0161406.
 PR 26-OCT-1999: 99US-0161359.
 PR 26-OCT-1999: 99US-0161360.
 PR 26-OCT-1999: 99US-0161361.
 PR 28-OCT-1999: 99US-0161920.
 PR 28-OCT-1999: 99US-0161922.
 PR 28-OCT-1999: 99US-0161993.
 PR 29-OCT-1999: 99US-0162142.

Query Match 10.9%; Score 113.4; DB 21; Length 1103;
 Best Local Similarity 50.6%; Pred. No. 8.3e-27;
 Matches 301; Conservative 0; Mismatches 291; Indels 3; Gaps 1;

QY 241 tataagggaatggtgacagcccttagggatcatatgaaggaaggcttctaagctt 300
 DB 345 tctcggaggttcttggaagcattgtgtacatagcaaggaagggttcttacta 404
 QY 301 tggcaagggaatgaccccgccattacagacgtatgtatcttcagggttcgaatg 360
 DB 405 tggaaagggttcttactctgcatcgtcgaatgtcctatttggaaggtccttagatgga 464
 QY 361 acatatgaacatctccgagaggtgtgtgttgccaaagtgaaatgacattccctt 420
 DB 465 atgtatgagcgggtgaanaactgtatgttggaagaagctttagtggatggtccatg 524
 QY 421 tggaaatcaatcttggaaggaatgagtggtgtgttccaggtttagccaatcca 480
 DB 525 agcaagaanaattcttgcgtgttgacacacagtgctggtatcatgtlagaacaatccc 584
 QY 481 actgaacctagtgaaaggttcagatgcaaatggaaggaanaaggaactggaaggaaccca 540
 DB 585 actgaaccttggaaggtttaggtctcaggccgaaggaanaa---ctagctcagggtgcgcca 641
 QY 541 ttgcgatttcgtgtgtacatcalttgcaaaaactttagctggaaggaataatcga 600
 DB 642 agacggaattcttgagcgctgaatgtcttacaacttltggaacaggaaggaatcga 701
 QY 601 gggccttggcagcgtggtggaaccaataataagagcagcactgtggaataatggagat 660
 DB 702 gctcttggagctgcttggcctaagctagcaagaacgtattatcaatgtcgtcgaa 761
 QY 661 ttaccacctatagatacaggaacactacttgtagttagatatacaccacttggaacat 720
 DB 762 tttagcgatgaatacaggaaggaagacatcttgaaggttcgggttccaatgacaac 821
 QY 721 ataatgaatcaggttatcaagtltatgttctgactgtagcttattcttggagaca 780
 DB 822 gttgtcacacatatctatctgactggtgagcaggaattcttctgtcgtcatcgctct 881
 QY 781 ccagcgagatgtcatcaanaagcagaataatgaatcaaccacagagataaacaaggaa 835
 DB 882 cctgtcagcgtgtgtaagtaagaataatgaggaattctgtgtcttacaaggaca 936

RESULT 11
 ID C33096
 ID C33096 standard; DNA; 1227 BP.
 AC C33096;
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1765.
 XX Hybridisation assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway;
 KM metabolic pathway; promoter; termination sequence; ss.
 XX Arabidopsis thaliana.
 OS
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PE 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999: 99US-0121825.
 PR 05-MAR-1999: 99US-0123180.
 PR 09-MAR-1999: 99US-0123548.
 PR 23-MAR-1999: 99US-0125788.
 PR 25-MAR-1999: 99US-0126264.
 PR 29-MAR-1999: 99US-0126785.
 PR 01-APR-1999: 99US-0127462.
 PR 06-APR-1999: 99US-0128234.
 PR 08-APR-1999: 99US-0128714.
 PR 16-APR-1999: 99US-0129845.
 PR 19-APR-1999: 99US-0130077.
 PR 21-APR-1999: 99US-0130449.
 PR 23-APR-1999: 99US-0130510.
 PR 28-APR-1999: 99US-0131449.
 PR 30-APR-1999: 99US-0132048.
 PR 30-APR-1999: 99US-0132407.
 PR 04-MAY-1999: 99US-0132484.
 PR 05-MAY-1999: 99US-0132485.
 PR 06-MAY-1999: 99US-0132486.
 PR 06-MAY-1999: 99US-0132487.
 PR 11-MAY-1999: 99US-0132863.
 PR 14-MAY-1999: 99US-0134256.
 PR 14-MAY-1999: 99US-0134218.
 PR 14-MAY-1999: 99US-0134219.
 PR 14-MAY-1999: 99US-0134221.
 PR 18-MAY-1999: 99US-0134370.
 PR 18-MAY-1999: 99US-0134768.
 PR 19-MAY-1999: 99US-0134941.
 PR 20-MAY-1999: 99US-0135124.
 PR 21-MAY-1999: 99US-0135353.
 PR 24-MAY-1999: 99US-0135629.
 PR 25-MAY-1999: 99US-0136021.
 PR 27-MAY-1999: 99US-0136392.
 PR 28-MAY-1999: 99US-0136782.
 PR 01-JUN-1999: 99US-0137222.
 PR 03-JUN-1999: 99US-0137528.
 PR 04-JUN-1999: 99US-0137502.
 PR 07-JUN-1999: 99US-0137724.
 PR 08-JUN-1999: 99US-0138094.
 PR 10-JUN-1999: 99US-0138540.
 PR 10-JUN-1999: 99US-0138847.
 PR 14-JUN-1999: 99US-0139119.
 PR 16-JUN-1999: 99US-0139452.
 PR 16-JUN-1999: 99US-0139453.
 PR 17-JUN-1999: 99US-0139492.
 PR 18-JUN-1999: 99US-0139454.
 PR 18-JUN-1999: 99US-0139455.
 PR 18-JUN-1999: 99US-0139456.
 PR 18-JUN-1999: 99US-0139457.
 PR 18-JUN-1999: 99US-0139458.

```
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144088.
PR 16-JUL-1999; 99US-0144088.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153738.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.2%; Score 106; DB 21; Length 1227;
Best Local Similarity 50.4%; Pred. No. 2.4e-24;
Matches 287; Conservative 0; Mismatches 280; Indels 3; Gaps 1;
```

PR	11-MAY-1999;	9905-013424256
PR	14-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134419
PR	14-MAY-1999;	9905-01343221
PR	14-MAY-1999;	9905-0134370
PR	18-MAY-1999;	9905-0134768
PR	19-MAY-1999;	9905-0134941
PR	20-MAY-1999;	9905-0135124
PR	21-MAY-1999;	9905-0135553
PR	24-MAY-1999;	9905-0135629
PR	25-MAY-1999;	9905-0136021
PR	27-MAY-1999;	9905-0136592
PR	28-MAY-1999;	9905-0136782
PR	01-JUN-1999;	9905-0137222
PR	03-JUN-1999;	9905-0137528
PR	04-JUN-1999;	9905-0137502
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138540
PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139119
PR	16-JUN-1999;	9905-0139452
PR	16-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139452
PR	18-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139457
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139763
PR	21-JUN-1999;	9905-0139617
PR	22-JUN-1999;	9905-0139899
PR	22-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	29-JUN-1999;	9905-0140991
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142390
PR	08-JUL-1999;	9905-0142803
PR	09-JUL-1999;	9905-0142920
PR	12-JUL-1999;	9905-0143277
PR	13-JUL-1999;	9905-0143542
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144086
PR	16-JUL-1999;	9905-0144325
PR	19-JUL-1999;	9905-0144321
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144334
PR	19-JUL-1999;	9905-0144335
PR	20-JUL-1999;	9905-0144352
PR	20-JUL-1999;	9905-0144632
PR	20-JUL-1999;	9905-0144684
PR	21-JUL-1999;	9905-0144814
PR	21-JUL-1999;	9905-0145086
PR	21-JUL-1999;	9905-0145088
PR	22-JUL-1999;	9905-0145085
PR	22-JUL-1999;	9905-0145087
PR	22-JUL-1999;	9905-0145089
PR	22-JUL-1999;	9905-0145192
PR	23-JUL-1999;	9905-0145195

XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

```

PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156566.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 23-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

Query Match 9.4%; Score 97.2; DB 21; Length 957;
 Best Local Similarity 49.6%; Pred. No. 1.6e-21;
 Matches 277; Conservative 0; Mismatches 278; Indels 3; Gaps 1;

```

QY 236 ccccatatgggaatggtgcgcacagccctaggatcatgaagagagagcttctaa 295
    || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 164 ccaagtataagagatcaatgttactctcgtacacatagctagagagaagatttcag 223
QY 236 agccttgcaagagatgacacccgcatttacagacagtagtcttctgaggtcgaa 355
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 224 gctcttggaagagtgcttattgtcagagcttcatcgcacatgtatctgtgttaaga 283
QY 356 tggtaacatataacatcccgagaggtgtgtttggcaaaagtgaaagtgcattacc 415
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 284 ttgggttaataagagcttgaaagacacttgtgtgaaagtacttattgtgcgataatc 343
QY 416 ccccttggaacacagtcattggaggatgatgtgtgttatttggccagttttagca 475
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 344 ctttatatacaaaagattcttcagcttcttgaactgagagctatagatatattgtagca 403
QY 476 atccaactgacctagtgaaaggttcagatgcaaatgaaagaaagaaactgtgaagaa 535
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 404 atccaactgaccttcttgaagctgtcagcttcatcaccattgtgaagctgtgaagaa 460
QY 536 aacccttgagattctgtgtgtacatcatgacttgcgaanaactltagctgaaggagaa 595
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 ttccctagagcgttatgagagagctgtgagagcttattccaccattgtgaagctgtgaag 520
QY 536 taccagagcgttgagagagctgtgagagcgttccaatatacaagaagagcagcagtgatag 655
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 ttatgagcgtatgagagcgttactgtgtcccaatatgtgccggaatgtctattgttaagtctg 580

```

```

QY 656 gagatttaaccactatgatatacagtgaaacactactgtgtattgaatacaccactgag 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 cagagtgtagctgtgtatgtatcaataaagagagaacatttgaaatctcgtcttcagag 640
QY 716 acaataatcagctcaagcgttatacaagttatgtctgtgactggtatctatctg 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 acagtgtttaactcatcactacagctgtgttagctgcaagctctctcgtctgcatcg 700
QY 776 gaaccaccagcagatgca 793
DB 701 gtcttccaatgagtcga 718

RESULT 14
A37719
ID A37719 standard; cDNA; 876 BP.
XX
AC A37719;
XX
XX 22-NOV-2000 (first entry)
DT
XX
DE Novel human protein coding sequence #1.
XX
XX Novel human protein; NHP; uncoupling protein; diagnosis; therapy;
KW metabolism; thermogenesis; obesity; cachexia; ss.
XX
XX Homo sapiens.
FH
FT Key Location/Qualifiers
FT CDS 1..876
FT /*tag= a
FT /product= NHP#1
XX
XX MO200047617-A1.
PD
XX 17-AUG-2000.
PE
XX 09-FEB-2000; 2000MO-US03425.
PF
XX
XX 09-FEB-1999; 9905-0119228.
PR
XX 08-OCT-1999; 9905-0158458.
XX
XX (LEXI-) LEXICON GENETICS INC.
PA
XX Turner CA, Mathur B, Zambrowicz B, Sands AT;
PI
XX WPI. 2000-558191/51.
DR
XX P-PSDB; Y90322.
XX
XX Nucleic acids encoding human uncoupling proteins useful for the study
XX and modulation of conditions such as obesity and cachexia -
XX
XX Claim 1: Page 25; 31pp; English.
XX
XX This sequence encodes a novel human protein (NHP) of the invention. The
XX NHP's of the invention are uncoupling proteins. The DNA and the encoded
XX protein may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate expression of the protein. The DNA
XX may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of NHP by
XX expressing inactive proteins or to supplement the patients own production
XX of NHP polypeptides. Additionally, the DNA may be used to produce the
XX protein. Conversely, complementary sequences and antisense nucleic acid
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own NHP genes and preventing their expression. The
XX DNA and the antisense molecules may also be used as DNA probes in
XX diagnostic assays to detect and quantitate the presence of similar
XX nucleic acid sequences in samples, and hence which patients may be in
XX need of restorative therapy. They may also be used to study the
XX expression and function of the NHP's and their role in metabolism by the
XX creation of transgenic and/or gene knock-out animals. The NHP
XX polypeptides may be used as antigens in the production of antibodies

```


Query Match 8.6%; Score 89.6; DB 21; Length 1031;
 Best Local Similarity 48.2%; Pred. No. 5.2e-19;
 Matches 417; Conservative 0; Mismatches 404; Indels 45; Gaps 4;

QY	117	cggtcgcgcgcgtacgcgtgcccgaagtaacaaccccttcctgagatccacaacacg	176
DB	168	cggtccttgcctcctatgttgcgcaggttcgcaccttcccttgatcttaacacacg	227
QY	177	actccaatgcgaaggaagcagcctcttgcctgcgttggagacggtgcagagaatctgc	236
DB	228	gcttgcagagcccaagccagatcgtatgttgcctcaagaataaa-----	276
QY	237	cccctaataggggaatgtgtgcacagccctagggatcatggaaggaaggttctctaa	296
DB	277	----tatagagagatgttctatgcctgttccgaatctataagaagaaggtctctg	332
QY	297	gcttggcgaagagtgacacccgcattacagacagtgatgtatctcggaggtcgaat	356
DB	333	ctgtatctcaggaatgcccctgttactaaagacagcatcatatgacccaataat	392
QY	357	ggtcacatatgacatctccgagaggtgtgttggcaaaagtgaagatgagcatatcc	416
DB	393	tgttatcttgaagagcttgaagcatatcttgaacgtttggaagatgagactctct	452
QY	417	ccttggaaatcagtcattggaagggaatgagctgtgttattgtgccaattttaagcaa	476
DB	453	aattaca-----tgatctgtgggttagtgcagagagtgatctcctactatctgcaa	506
QY	477	tccaactgacctagtgaaaggttcagatgcaatgtaaggaagaaagaaacttgaaagaa	536
DB	507	tcccaactgactgtcttaaaagttcgaatgcaggtcgaaggaagttgttccaaggagcat	566
QY	537	accattgcgattctgtgtgtacatcatgtcatttgcacaaatcctagctgaagaggaat	596
DB	567	gattggcagcttcatatgacatatacca-----gcaagaaggtac	605
QY	597	acgaagggttctggcagagctgtgtacccaatatacaagaagcagacgttgaataggg	656
DB	606	caggggctcgtgaggggtgtgttcccaactgtcagcgtgcgcaatcgttgtggagt	665
QY	657	agatttaaccacttaatacagtgaaacactactgtgatatgaatacaccacttgagga	716
DB	666	agaagctgcccgttattgacattacccaagaagacaccctgatttccagaaatgtcggaga	725
QY	717	caatatcagctcagcgttatacaagttatgttctgacgtgagcttctatctcggg	776
DB	726	caaattttaacacacttgttccagtttcaacctgtgtgttgcctgggctctgcacac	785
QY	777	aacacagcgcgaatgtatcaaaaagcaaaataatgaatacacaagataaacaagaag	836
DB	786	taacctgtgagtgtgtgagaaactcgaatgtatgaatca---gaggcaatagtgagaca	842
QY	837	gggaacttctgtataatcactgactgactgtgattcagcgtgttcaaggtgaagatt	896
DB	843	tgtgaacctctacaaggtacttgtatgttatttaagatgtgaaagcatgagagatt	902
QY	897	catgagctcctataaagcgttttaccatctggtcgtgagaatgacccctgttcaatgt	956
DB	903	tttgcactctataaagcgttctggccaactggtcgtgacttgacccctggaacatcat	962
QY	957	gttctgcttactatgaaanaatca	982
DB	963	ttttttattacctatgacagctca	988

QY 623 ccaataaagaagcagcactgltgaatataatggagatttaaccacttaatgacagltga 682
DB 778 CCAACATACAGAAGAAATGCCATTGTCACTGCTGAGATGCTGACCTACGACATCATCA 837
QY 683 aacactactgttatgtaataacacacttgagacaaatcatatgactacaggtttacaa 742
DB 838 AGAGAAAGTTGCTGAGAGTCTCACTGTTACTGACAACTTCCCTGCTACTTGTCTCTG 897
QY 743 gtttaagcttgagactgagacttctatcttggaacacacagccagatgcatacaaaaga 802
DB 898 CTTTGAGACTGCTTCTGTGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 957
QY 803 gaataatgaataacacagagataaacaagaaggagactttgtataaatacatgactg 862
DB 958 GATACATGAACCTCCCTAGGACAGGTACCGACCCCTCTGC----- 999
QY 863 actgcttgatcaagctgttcaagtgagagattcatatgactataaaggcttttac 922
DB 1000 ACTGATGCTGAAGATGCTGCTCAAGAGGACCCAGCTTCTCAAAAGATTGTGTC 1059
QY 923 catcttgctgagaaatgacccttgtaagtgtgtctgacttaataaaggcttttac 982
DB 1060 CCTCTTCTGCTGCTGCTGAGAGCTTGAACGTATGATGTTGTACATATGAGCAACTCA 1119

RESULT 2

US-09-172-528-3

Sequence 3, Application US/09172528

Patent No. 5952469

GENERAL INFORMATION:

APPLICANT: Zhang, Ning

APPLICANT: Amaral, M. Catherine

APPLICANT: Chen, Jin-Long

TITLE OF INVENTION: UCP3 Genes

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/172,528

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/937,466

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: T97-009

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1949 base pairs

TYPE: nucleic acid

STRADEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-172-528-3

Query Match 4.9%; Score 50.8; DB 2; Length 1949;

Best Local Similarity 47.6%; Pred. No. 2.5e-07;
Matches 200; Conservative 0; Mismatches 202; Indels 18; Gaps 1;
QY 563 atgcattgcaaaaatttagctgaaggaggaatcaggaaggtttggagagcttggatc 622
DB 718 ATGCTACAGAACCATCGCCAGGAGAGAGATCAAGGGCTGTGGAAGGACTTGCC 777
QY 623 ccaataaagaagcagcactgltgaatataatggagatttaaccacttaatgacagltga 682
DB 778 CCAACATACAGAAGAAATGCCATTGTCACTGCTGAGATGCTGACCTACGACATCATCA 837
QY 683 aacactactgttatgtaataacacacttgagacaaatcatatgactacaggtttacaa 742
DB 838 AGAGAAAGTTGCTGAGAGTCTCACTGTTACTGACAACTTCCCTGCTACTTGTCTCTG 897
QY 743 gtttaagcttgagactgagacttctatcttggaacacacagccagatgcatacaaaaga 802
DB 898 CTTTGAGACTGCTTCTGTGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 957
QY 803 gaataatgaataacacagagataaacaagaaggagactttgtataaatacatgactg 862
DB 958 GATACATGAACCTCCCTAGGACAGGTACCGACCCCTCTGC----- 999
QY 863 actgcttgatcaagctgttcaagtgagagattcatatgactataaaggcttttac 922
DB 1000 ACTGATGCTGAAGATGCTGCTCAAGAGGACCCAGCTTCTCAAAAGATTGTGTC 1059
QY 923 catcttgctgagaaatgacccttgtaagtgtgtctgacttaataaaggcttttac 982
DB 1060 CCTCTTCTGCTGCTGCTGAGAGCTTGAACGTATGATGTTGTACATATGAGCAACTCA 1119

RESULT 3

US-09-318-199-3

Sequence 3, Application US/09318199

Patent No. 6025469

GENERAL INFORMATION:

APPLICANT: Zhang, Ning

APPLICANT: Amaral, M. Catherine

APPLICANT: Chen, Jin-Long

TITLE OF INVENTION: UCP3 Genes

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/318,199

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/937,466

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: T97-009

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1949 base pairs

TYPE: nucleic acid

STRADEDNESS: double

US-09-318-199-3

TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-318-199-3

Query Match 4.9%; Score 50.8; DB 3; Length 1949;
Best Local Similarity 47.6%; Pred. No. 2.5e-07;
Matches 200; Conservative 0; Mismatches 202; Indels 18; Gaps 1;

QY 563 atgatttgcacaaattccttagctgaaggaagaaacagaggtcttgagcagcttggtac 622
DB 718 ATGCTTCAGAAACCATGCGCCAGGAGAGAGAGATCAGGGGCTGTGGAAGGACTTGGC 777
QY 623 caatatcaaaagacagcagcttgatgatatgagagatttaaccattatatacagta 682
DB 778 CCACATCACAAGAAATGCCATTGTCAACTGTCTGAGATGTTACCTACACATCA 837
QY 683 aacactacttgatltgaatacaccacttgaagacaatalcatagctacaggtttatca 742
DB 838 AGGAGAGTGTGCTGGAGTCTACCTGTTTACTGACAACTTCCCTGTCATTGTCTGTG 897
QY 743 gttatgttctgagctgtagcttctatcttggaacacacagccagatgltacaaagca 802
DB 898 CTTTGGAGCTGGCTTGTGTCCACAGTGTGGCTCCCGGTGGATGTGTTAAAGACCC 957
QY 803 gaataatgaatacaacacagagataaacaagaaggagacttltgataatacatcagctg 862
DB 958 GATCATGTAACCGTCCCTTAGGAGATACCCAGCCCTCTTC----- 999
QY 863 actgcttgatcagagctgttcaaggtgaagatcatgagctatataaagcttttac 922
DB 1000 ACTGTATGCTGAAGATGTGGCTCAGAGGAGACCCAGGCTTCTACAAAGATTTGTGTC 1059
QY 923 catcttgctgagagatgaaccttggtcaatggtgtctggtctacttaataaatac 982
DB 1060 CTCTCTTGTGCTGCTGTGGAGCTTGAACGATGATGTTTGAACATATGACAGACTGA 1119

RESULT 4

US-08-518-878B-38
Sequence 38, Application US/0851878B
Patent No. 5702902
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518,878B
FILING DATE: 23-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:

LENGTH: 1255 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-518-878B-38

Query Match 4.8%; Score 49.4; DB 1; Length 1255;
Best Local Similarity 49.1%; Pred. No. 5.7e-07;
Matches 199; Conservative 0; Mismatches 191; Indels 15; Gaps 2;

QY 103 aaatcctactgctcgcgctgctacacgctggcagcagctacgaacattcccttgat 162
DB 16 AAGTTTCTGGGGCTGGGACAGCTGCTGCATGCCAATGTCATCACTTCCCTGTGAT 75
QY 163 ctcaaaaactcgaactcacaatgaaggaagacagctctgctcggttggaagcgt 222
DB 76 ACTCTAAAGTCCGGTTACAGATCCAAAGAGAAAG-----TCAAGGGCCAGTCCGC 126
QY 223 gcaagagaatctgcccctataagggaatggtgcaacagccctaggaatcatggaag 282
DB 127 GCTACAGTCACGCGCCAGTACCGGCTGTGATGGGACACATTCGACCATGTCGTA 186
QY 283 gaagcttctcaagcttggcaaggaatgaacccgcatltaagaacgtagtat 342
DB 187 GAGGGCCCCGAGAGCTCTTACATGAGGCTGTGCTGCGGCTGACGCGCAATGAGCTT 246
QY 343 tctgagagtcgaatggtcaacatgaacatccgagaaggttggttgcaaaagtga 402
DB 247 GCCCTGTGCCCATACGCGCTGTATGATCTGTCAACA-----GTCTACACCAAGGCG 300
QY 403 gataagcatlcccttggaatcaatgaatcagcttggaaggaatgagctggttatgac 462
DB 301 TCTAGCATGCGCAGATTTGGAGCGCCCTCTTACAGGACACACAGGTGCTGCT 360
QY 463 cagttttagccaatccacagctagtggaaggttcagatgca 507
DB 361 GTGGCTGTGGCCAGCCAGCGATGTGTAAAGTCCGATTCCAA 405

RESULT 5

US-08-294-522B-38
Sequence 38, Application US/08294522B
Patent No. 5741666
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/294,522B
FILING DATE: 23-AUG-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:

FILING DATE: 23-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-807-861A-38

Query Match 4.8%; Score 49.4; DB 2; Length 1596;
Best Local Similarity 49.1%; Pred. No. 6.8e-07;
Matches 199; Conservative 0; Mismatches 191; Indels 15; Gaps 2;

QY 103 aaattcctactgctcggcgtcgccgctacccgtggccgagctagaacattcccttgat 162
DB 358 AAGTTTCTTGGGGCTGGACAGCTGCTGCATGCAGATCTCATCACTTCTCTGGAT 417
QY 163 ctcaaaaactcgtactccaatgcaagagagagcagctctgctcggttgaggagcgt 222
DB 418 ACTGCTTAAGTCCGGTTACAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
QY 223 gcaagagaatctgcccctctagagggaatggtgacagccctaggagatcatgaaag 282
DB 469 GCTACAGTACAGCCGCCAGTACCGCGGTGTATGAGGACACATTCTGACCATGTCGTA 528
QY 283 gaaggtcttctaaagcttggcaagagatgacacccgcatllacagacagctagtgat 342
DB 529 GAGGGCCCCGAAAGCTCTACAAATGGGCTGTTGCCGCTGCAGCCGCAAAATAGCTTT 588
QY 343 tctgagagtcgaatggtacacatgaaacatcccgagagagtggttggttgcaaaagt 402
DB 589 GCCCTCTGCGCATCGCGCTGTATGATTTCTGTCAACA-----GTTCTACACCAAGGCG 642
QY 403 gatgagattatccctcttgaaatcagtcattgagagagatgagtgctggtattggc 462
DB 643 TCTGAGCATGCCAGCATTTGGAGCGCCCTCTACAGGACACACAGAGTGCCCTGCT 702
QY 463 cagttttagccaatccaactgactagtgaaagttcagatgcaa 507
DB 703 GTGGCTGTGGCCCGCCACGAGATGTGTAAAGTCCGATTCCAA 747

RESULT 8
US-09-210-681-38
Sequence 38, Application US/09210681
Patent No. 6057109
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
REGULATION OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/807,861
FILING DATE: 26-FEB-1997
APPLICATION NUMBER: US 08/518,878
FILING DATE: 23-AUG-1995
APPLICATION NUMBER: US 08/470,868
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/294,522
FILING DATE: 23-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-210-681-38

Query Match 4.8%; Score 49.4; DB 3; Length 1596;
Best Local Similarity 49.1%; Pred. No. 6.8e-07;
Matches 199; Conservative 0; Mismatches 191; Indels 15; Gaps 2;

QY 103 aaattcctactgctcggcgtcgccgctacccgtggccgagctagaacattcccttgat 162
DB 358 AAGTTTCTTGGGGCTGGACAGCTGCTGCATGCAGATCTCATCACTTCTCTGGAT 417
QY 163 ctcaaaaactcgtactccaatgcaagagagagcagctctgctcggttgaggagcgt 222
DB 418 ACTGCTTAAGTCCGGTTACAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
QY 223 gcaagagaatctgcccctctagagggaatggtgacagccctaggagatcatgaaag 282
DB 469 GCTACAGTACAGCCGCCAGTACCGCGGTGTATGAGGACACATTCTGACCATGTCGTA 528
QY 283 gaaggtcttctaaagcttggcaagagatgacacccgcatllacagacagctagtgat 342
DB 529 GAGGGCCCCGAAAGCTCTACAAATGGGCTGTTGCCGCTGCAGCCGCAAAATAGCTTT 588
QY 343 tctgagagtcgaatggtacacatgaaacatcccgagagagtggttggttgcaaaagt 402
DB 589 GCCCTCTGCGCATCGCGCTGTATGATTTCTGTCAACA-----GTTCTACACCAAGGCG 642
QY 403 gatgagattatccctcttgaaatcagtcattgagagagatgagtgctggtattggc 462
DB 643 TCTGAGCATGCCAGCATTTGGAGCGCCCTCTACAGGACACACAGAGTGCCCTGCT 702
QY 463 cagttttagccaatccaactgactagtgaaagttcagatgcaa 507
DB 703 GTGGCTGTGGCCCGCCACGAGATGTGTAAAGTCCGATTCCAA 747

RESULT 9
US-08-946-719A-38
Sequence 38, Application US/08946719A
Patent No. 6121017
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 64

[illegible]

Db 838 AGGAGAAGTGTGAGACTTCACCTGTTTACTGACAACTTCCCTGTGACTTGTCTCTG 897
 QY 743 gttatgtcttgagactgtagcttctatctctggaacacccagcgatgcatcaaaagca 802
 Db 898 CCTTTGAGCTGGCTCTGTGTCACACAGTGGTGGCTCCCGGTGGATGTGTAAAGACC 957
 QY 803 gaataatgaatcaacacagagataaacaagaaggagacttggat 849
 Db 958 GATACATGAACGCTCCCTAGGCAAGTACCGACAGCACTCAGAAAT 1004

RESULT 15

US-09-318-199-5
 ; Sequence 5, Application US/09318199
 ; Patent No. 6025469
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang, Ning
 ; APPLICANT: Amaral, M. Catherine
 ; APPLICANT: Chen, Jin-Long
 ; TITLE OF INVENTION: UCP3 Genes
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 75 DENISE DRIVE
 ; CITY: HILLSBOROUGH
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94010
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/318,199
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/937,466
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OSMAN, RICHARD A
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: T97-009
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 343-4341
 ; TELEFAX: (650) 343-4342
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1777 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-09-318-199-5

Query Match 4.7%; Score 48.6; DB 3; Length 1777;

Best Local Similarity 48.1%; Pred. No. 1.4e-06;

Matches 138; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 563 atgcattgcacaaatcttagctgaagaggaataacgagggccttggcaagcgcgggtac 622
 Db 718 ATGCCCTACAGAACCATGCCAGGAGGAGAGAGTCAAGGGCCTGTGAAAGGAGACTTGGC 777
 QY 623 ccaatatacaagaagcagcactgtagaataatggagagtttaaccactatgatacaagta 682
 Db 778 CCAACATCACAGAAATGCCATTGTCACTGTGCTGAGATGTGACCTACGACATCATCA 837
 QY 683 aacactactgtagtataatgaataccacttagagacaatatcatgactacggtttatcaa 742
 Db 838 AGGAGAAGTGTGAGACTTCACCTGTTTACTGACAACTTCCCTGTGACTTGTCTCTG 897

QY 743 gttatgtcttgagactgtagcttctatctctggaacacccagcgatgcatcaaaagca 802
 Db 898 CCTTTGAGCTGGCTCTGTGTCACACAGTGGTGGCTCCCGGTGGATGTGTAAAGACC 957
 QY 803 gaataatgaatcaacacagagataaacaagaaggagacttggat 849
 Db 958 GATACATGAACGCTCCCTAGGCAAGTACCGACAGCACTCAGAAAT 1004

Search completed: April 28, 2001, 18:18:31
 Job time: 2730 sec

Trying 3106016892...Open

Welcome to STN International! Enter xxx
LOGINID:ssppa1816mt
PASSWORD
TERMINAL (ENTER 1, 2, 3, OR 7):2

***** Welcome to STN International *****

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America
NEWS 2 Dec 17 The CA Lexicon available in the CAPLUS and CA files
NEWS 3 Feb 06 Engineering Information Enccompass files have new names
NEWS 4 Feb 16 TOXLINE no longer being updated
NEWS 5 Apr 23 Search Devent WPINDEX by chemical structure
NEWS 6 Apr 23 PRE-1987 REFERENCES NOW SEARCHABLE IN CAPLUS
AND CA

NEWS EXPRESS April 18 CURRENT WINDOWS VERSION IS V6.0
CURRENT MACINTOSH VERSION IS V5.0C (ENG) AND V5.0AB (JP).
AND CURRENT DISCOVER FILE IS DATED 04/06

NEWS HOURS STN Operating Hours Plus Help Desk Availability
NEWS INTER General Internet Information
NEWS LOGIN Welcome Banner and News Items
NEWS PHONE Direct Dial and Telecommunication Network Access to STN
NEWS WWW CAS World Wide Web Site (General Information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

***** STN Columbus *****

FILE 'HOME' ENTERED AT 12:39:45 ON 03 MAY 2001

=> file medicine cancerlit biosis embase scisearch

COST IN U.S. DOLLARS	ENTRY	SESSION	SINCE FILE	TOTAL
FULL ESTIMATED COST	0.15	0.15		

FILE 'MEDLINE' ENTERED AT 12:40:09 ON 03 MAY 2001

FILE 'CANCERLIT' ENTERED AT 12:40:09 ON 03 MAY 2001

FILE 'BIOSIS' ENTERED AT 12:40:09 ON 03 MAY 2001

COPYRIGHT (C) 2001 BIOSIS(R)

FILE 'EMBASE' ENTERED AT 12:40:09 ON 03 MAY 2001
COPYRIGHT (C) 2001 Elsevier Science B.V. All rights reserved.

FILE 'SCISEARCH' ENTERED AT 12:40:09 ON 03 MAY 2001
COPYRIGHT (C) 2001 Institute for Scientific Information (ISI) (R)

=> s ucpr4

L1 29 UCPR4

=> dup rem

ENTER ,# LIST OR (END):11

PROCESSING COMPLETED FOR L1

L2 10 DUP REM L1 (19 DUPLICATES REMOVED)

=> d 12 1-10 bib ab

L2 ANSWER 1 OF 10 MEDLINE

DN 2001227289 MEDLINE

AN 21134422 PubMed ID: 11239488

TI Homologues of the uncoupling protein from brown adipose tissue (UCP1):

UCP2, UCP3, BMCP1 and UCP4.

AU Bouillaud F, Couplan E, Pecqueur C, Riquier D
CS CEREMOD, C.N.R.S., UPR 9078, 9 rue Jules Hetzel, 82190 Meudon,
France.

bouillaud@intbiogen.fr
SO BIOCHIMICA ET BIOPHYSICA ACTA, (2001 Mar 1) 1504 (1) 107-19. Ref: 83
Journal code: AOW; 0271513. ISSN: 0006-3002.

CY Netherlands

DT Journal; Article: (JOURNAL ARTICLE)
(REVIEW, TUTORIAL)

LA English

FS Priority Journals

EM 200104

EW Entered STN: 20010502

Last Updated on STN: 20010502

Entered PubMed: 20010312

Entered Medline: 20010426

L2 ANSWER 2 OF 10 MEDLINE

AN 2000412260 MEDLINE

DN 20387222 PubMed ID: 10928996

TI Characterization of novel UCP5/BMCP1 isoforms and differential regulation of UCP4 and UCP5 expression through dietary or temperature manipulation.

AU Yu X X, Mao W, Zhong A, Schow P, Brush J, Sherwood S W, Adams S H; Pan G
CS Departments of Endocrinology, Molecular Biology, and Bioassay and Biomege, Genentech, Inc., South San Francisco, California 94080, USA.

SO FASEB JOURNAL, (2000 Aug) 14 (11) 1611-8.
Journal code: FAS; 8804484. ISSN: 0892-6633.

CY United States

DT Journal; Article: (JOURNAL ARTICLE)

LA English

FS Priority Journals

EM 200008

EW Entered STN: 20000907

Last Updated on STN: 20000907

Entered Medline: 20000825

AB Mitochondrial uncoupling proteins have been implicated in the maintenance of metabolic rate and adaptational thermoregulation. We recently reported the identification of a brain-specific mitochondrial uncoupling protein homologue, UCP4. Here we characterized another newly described member of the uncoupling protein family, termed UCP5 (also called BMCP1). UCP5 transcripts are present in multiple human and mouse tissues, with an especially high abundance in the brain and testis. Expression of UCP5 in mammalian cells reduces the mitochondrial membrane potential. Multiple isoforms of UCP5 were identified and exhibited tissue-specific distribution and different potency in reduction of membrane potential. Furthermore, the mRNA abundance of both UCP4 and UCP5 is modulated by nutritional status or temperature in a tissue-specific manner in mice. Brain UCP4 and UCP5 mRNA transcripts rose by 1.5- and 1.7-fold, respectively, and liver UCP5 expression increased by 1.8-fold in response to acute cold exposure. A high-fat diet increased UCP5 mRNA in liver by 1.6-fold selectively in the obesity-resistant AU but not in the obesity-prone C57BL/6J mouse strain. Liver UCP5 expression decreased significantly with a 24 h fast and was restored to the normal level after refeeding. In contrast, brain transcripts for both genes were not significantly altered by fasting or high-fat diet. These findings are consistent with the notion that UCP4 and UCP5 may be involved in tissue-specific thermoregulation and metabolic changes associated with nutritional status.

L2 ANSWER 3 OF 10 MEDLINE

AN 2000202381 MEDLINE

DN 20202381 PubMed ID: 10736318

TI Uncoupling protein homologs: emerging views of physiological function.

AU Adams S H

CS Department of Endocrinology, Genentech, Incorporated, South San Francisco.

CA 94080, USA.

SO JOURNAL OF NUTRITION, (2000 Apr) 130 (4) 711-4. Ref: 57

Journal code: JEV; 0404243. ISSN: 0022-3166

CY United States

DT Journal; Article: (JOURNAL ARTICLE)

General Review, (REVIEW)

(REVIEW, TUTORIAL)

LA English

FS Priority Journals

EM 200004

EW Entered STN: 20000505

Last Updated on STN: 20000505

Entered Medline: 20000424

AB The widespread occurrence of excess weight and related diseases demands that efforts be made to understand energy expenditure from the gene to the whole animal. For some time, it has been understood that mitochondrial oxidation of fuels generates an electrochemical gradient via outward pumping of protons by the electron transport chain. ATP production via F(1)F(0) ATP synthase is then facilitated by the inward flux of protons down the gradient. There is a growing appreciation that a significant portion of the metabolic rate of endotherms is attributable to counteracting "proton leak" (uncoupling), wherein a flux of protons down the electrochemical gradient generates heat independently of ATP production. Proton leak is especially apparent in thermogenic brown adipose tissue, which expresses a tissue-specific uncoupling protein (UCP1). The recent discovery of widely expressed putative UCP1 homologs [UCP2, UCP3, UCP4, UCP5] brain mitochondrial carrier protein-1 (BMCP1)] raised the possibility that innate proton leak and metabolic rate are regulated by UCP1-like proteins. On the basis of current published data, one may not exclude the possibility that UCP homologs influence metabolic rate.

L2 ANSWER 4 OF 10 MEDLINE

AN 2000233287 MEDLINE

DN 20233287 PubMed ID: 10772343

TI Specific sequence of motifs of mitochondrial uncoupling proteins.

AU Jezek P, Urbankova E

CS Department of Membrane Transport Biophysics, Institute of Physiology, Academy of Sciences of the Czech Republic, Prague.

jezek@sun1.biomed.cas.cz

SO JUBMB Life, (2000 Jan 49 (1) 63-70

Journal code: DJT; 10088706. ISSN: 1521-6543.

CY ENGLAND: United Kingdom

DT Journal; Article: (JOURNAL ARTICLE)

LA English

FS Priority Journals

EM 200006

EW Entered STN: 20000706

Last Updated on STN: 20000706

Entered Medline: 20000626

AB We have searched for the exclusivity of common sequence motifs of the mitochondrial uncoupling proteins (UCP1, UCP2, UCP3, UCP4, BMCP1, and plant UCP [PUMF]) within the gene family of mitochondrial anion carrier proteins. The UCP-specific sequences, "UCP signatures", were found in the first, second, and fourth alpha-helices. First: Ala/Ser-Cys/Thr/n-Phe-Ala/Gly (negatively charged residue)/n-Phe-n-Cys-Thr-Phe/n, second: Gly/Ala-Ileu/Glu-Gln-X (positively charged residue)/n-H-n-Cys-Ser/Pro/n-X-n-Ser-OH/Gly-n (positively charged residue)/Ileu/Met-Gly/Val-n/Thr, fourth: Pro-Asn/Thr-n-X (positively charged residue)/Ileu/Met-Gly/Val-n/Thr. Pro-Asn/Thr-n-X (positively charged residue)/Asn/Ser/Ala-n-Ileu/n-Asn/Val-Cys/n-n/Thr (negatively charged residue)/n-n/Thr/Pro-OH/Val (n, nonpolar, pH, aromatic; positively charged residue/negatively charged residue, charged residue). The second and part of the third signature are also present in the yeast dicarboxylate transporter. The UCP signature excluding BMCP1 was also found in the second matrix segment, [positively charged residue]/Pro/Ileu/Ileu/Ileu (positively charged residue)/pH-X-Gly/Ser-Thr/n-X-NH (negatively charged residue)/Ala-pH. These UCP signatures are thought to be involved in fatty acid anion binding and translocation.

L2 ANSWER 5 OF 10 BIOSIS COPYRIGHT 2001 BIOSIS

AN 2001 86714 BIOSIS

DN PREV200100088714

TI BMCP1: a neuronal protein which regulates mitochondrial free radical production and respiratory rate.

AU Kim-Han, J. S. (1); Reichert, S. A.; Dugan, L. L.

CS (1) Washington University School of Medicine, Saint Louis, MO USA

SO Society for Neuroscience Abstracts, (2000) Vol. 26, No. 1-2, pp Abstract No. 378.16, print.

Meeting Info.: 30th Annual Meeting of the Society of Neuroscience

Orleans, LA, USA November 04-08, 2000 Society for Neuroscience

ISSN: 0190-5295.

DT Conference

LA English

SL English

AB Mitochondrial uncoupling proteins (UCPs) have been identified in various tissues, including brain. UCPs were first identified in brown adipose tissue, where they are thermogenic by uncoupling mitochondrial respiration from ATP production. Brain mitochondrial carrier protein 1 (BMCP1) is related to the UCPs, showing protein sequence homology of 34, 39, 39, and 39% with UCP1, UCP2, UCP3, and UCP4, respectively. We examined the function and characteristics of BMCP1 in brain using cortical neuronal cultures, and mouse and human brain tissue. We raised rabbit polyclonal antibody against the 14 amino acids corresponding to the protein sequence between the 1st and 2nd transmembrane domains of mouse BMCP1. This

AB

identified a single band of 36 kDa on western blot analysis of both adult mouse and human cortex, with lower levels of expression in mouse heart and kidney, and no expression in liver or lung. BMCP1 was expressed only in neurons, but not astrocytes, and protein expression in both cortical neurons and mouse brain was developmentally upregulated. BMCP1 was localized to mitochondria, as shown by immunocytochemistry and western blot analysis. Overexpression of BMCP1 in GT1-1 cells resulted in a significant reduction of basal superoxide production determined by dihydroethidium oxidation. Oxygen consumption rates using succinate as substrate were not affected by overexpression of BMCP1. However, oxygen consumption by EC-A was more profound in BMCP1 cells than in control cells. Because chronic overexpression of BMCP1 appeared to affect the phenotype and proliferation of GT1-1 cells, we began to study the effect of induced BMCP1 overexpression using tetracycline-inducible PC12 transfectants.

L2 ANSWER 6 OF 10 MEDLINE

AN 2000117196 MEDLINE

DN 20117196 PubMed ID: 1063471

TI Uncoupling protein-a useful energy dissipator.

AU Klingenberg M

CS Institute of Physical Biochemistry, University of Munich, Germany.

SO JOURNAL OF BIOENERGETICS AND BIOMEMBRANES. (1998 Oct) 31

(5) 419-30. Ref:

80

Journal code: HIC; 7701859. ISSN: 0145-479X.

CY United States

DT Journal; Article. (JOURNAL ARTICLE)

General Review. (REVIEW)

(REVIEW, TUTORIAL)

LA English

FS Priority Journals

EM 200003

ED Entered STN: 20000320

Last Updated on STN: 20000320

Entered Medline: 20000306

AB The structure/function relationship in the uncoupling proteins (UCP) is reviewed, stressing UCP from brown adipose tissue (UCP1) since, so far, nearly no biochemistry is known for the UCP variants UCP2, UCP3, and UCP4. The transport for H⁺ and Cl⁻ and its dependence on fatty acids in reconstituted vesicles is described. The inhibition and binding of nucleotides to UCP1, in particular, the pH dependence and two-stage binding are analyzed. A model for the role of fatty acid in H⁺ transport is shown. The role of specific residues in UCP1 is analyzed by directed mutagenesis in a yeast expression system. The different regulation by the cellular energy potential of UCP1 versus UCP3 is discussed.

L2 ANSWER 7 OF 10 MEDLINE

AN 1989148824 MEDLINE

DN 99148824 PubMed ID: 10025957

TI UCP4, a novel brain-specific mitochondrial protein that reduces

membrane potential in mammalian cells.

CM Erratum in: FEBS Lett 1989 Apr 23;449(2-3):283

AU Mao W, Yu X X, Zhong A, Li W, Brush J, Sherwood S W, Adams S H, Pan G

CS Department of Endocrinology, Genentech, Inc., South San Francisco, CA

94080. USA.

SO FEBS LETTERS. (1989 Jan 28) 443 (3) 326-30.

Journal code: EUH; 0155157. ISSN: 0014-5793.

CY Netherlands

DT Journal; Article. (JOURNAL ARTICLE)

FS Priority Journals

OS GENBANK:AF110532

EM 199903

ED Entered STN: 19990324

Last Updated on STN: 20000303

Entered Medline: 19990311

AB Uncoupling proteins (UCPs) are a family of mitochondrial transporter proteins that have been implicated in thermoregulatory heat production and maintenance of the basal metabolic rate. We have identified and partially characterized a novel member of the human uncoupling protein family, termed uncoupling protein-4 (UCP4). Protein sequence analyses showed that UCP4 is most related to UCP3 and possesses features characteristic of mitochondrial transporter proteins. Unlike other known UCPs, UCP4 transcripts are exclusively expressed in both fetal and adult brain tissues. UCP4 maps to human chromosome 9p11.2-q12. Consistent with its potential role as an uncoupling protein, UCP4 is localized to the mitochondria and its ectopic expression in mammalian cells reduces mitochondrial membrane potential. These findings suggest that UCP4 may be involved in thermoregulatory heat production and metabolism in the brain.

L2 ANSWER 8 OF 10 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.

AN 1999146875 EMBASE

TI Erratum: UCP4, a novel brain-specific mitochondrial protein that

reduces membrane potential in mammalian cells (FEBS Letters 443 (1999)

(326-330) PII: S001457939801713X)

AU Mao W, Yu X X, Zhong A, Li W, Brush J, Sherwood S W, Adams S H, Pan G

CS G

CS G, Pan, Department of Endocrinology, Genentech, Inc., 1 DNA Way, South

San

Francisco, CA 94080, United States, jippan@gene.com

SO FEBS Letters. (1999) 449(2-3) (293).

Ref: 0

ISSN: 0014-5793. CODEN: FEBLAL

PUI S 0014-5793(99)00318-X

CY Netherlands

DT Journal; Errata

FS 029 Clinical Biochemistry

LA English

L2 ANSWER 9 OF 10 SCISEARCH COPYRIGHT 2001 ISI (R)

AN 1999-360327 SCISEARCH

GA The Genuine Article (R) Number: 192LW

TI UCP4, a novel brain-specific mitochondrial protein that reduces

membrane potential in mammalian cells (vol 443, pg 326, 1999)

AU Mao W G, Yu X X, Zhong A, Li W L, Brush J, Sherwood S W, Adams S H, Pan G

H (Reprint)

CS GENENTECH INC, DEPT ENDOCRINOL, M-S-371, 1 DNA WAY, S SAN

FRANCISCO, CA

94080 (Reprint); GENENTECH INC, DEPT ENDOCRINOL, S SAN

FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA

94080; GENENTECH INC, DEPT MOL BIOL, S SAN FRANCISCO, CA



General Review, (REVIEW)
(REVIEW, TUTORIAL)

LA English
FS Priority Journals
EM 200104
ED Entered STN: 20010502
Last Updated on STN: 20010502
Entered PubMed 20010312
Entered Medline: 20010426

L4 ANSWER 3 OF 8 MEDLINE
AN 2000412260 MEDLINE
DN 20387222 PubMed ID: 10928996
TI Characterization of novel UCP5/BMCP1 isoforms and differential regulation of UCP4 and UCP5 expression through dietary or temperature manipulation, AU Yu X X; Mao W; Zhong A; Schow P; Brush J; Sherwood S W; Adams S H; Pan G
CS Departments of Endocrinology, Molecular Biology, and Bioassay and Biomeq, Genentech, Inc., South San Francisco, California 94080, USA.
SO FASEB JOURNAL, (2000 Aug) 14 (11) 1611-8.
CY Journal code: FAS, 8804484, ISSN: 0892-6638.
CJ United States
DT Journal, Article, (JOURNAL ARTICLE)
English
FS Priority Journals
EM 200008
ED Entered STN: 20000907
Last Updated on STN: 20000907
Entered Medline: 20000825

AB Mitochondrial uncoupling proteins have been implicated in the maintenance of metabolic rate and adaptational thermoregulation. We recently reported the identification of a brain-specific mitochondrial uncoupling protein homologue, UCP4, here we characterized another newly described member

of the uncoupling protein family, termed UCP5 (also called BMCP1). UCP5 transcripts are present in multiple human and mouse tissues, with an especially high abundance in the brain and testis. Expression of UCP5 in mammalian cells reduces the mitochondrial membrane potential. Multiple isoforms of UCP5 were identified and exhibited tissue-specific distribution and different potency in reduction of membrane potential. Furthermore, the mRNA abundance of both UCP4 and UCP5 is modulated by

nutritional status or temperature in a tissue-specific manner in mice. Brain UCP4 and UCP5 mRNA transcripts rose by 1.5- and 1.7-fold, respectively, and liver UCP5 expression increased by 1.8-fold in response to acute cold exposure. A high-fat diet increased UCP5 mRNA in liver by 1.6-fold selectively in the obesity-resistant AU but not in the obesity-prone C57BL/6J mouse strain. Liver UCP5 expression decreased significantly with a 24 h fast and was restored to the normal level after refeeding. In contrast, brain transcripts for both genes were not significantly altered by fasting or high-fat diet. These findings are consistent with the notion that UCP4 and UCP5 may be involved in tissue-specific thermoregulation and metabolic changes associated with nutritional status.

L4 ANSWER 4 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS

AN 2000197467 BIOSIS
DN PRE200000187467
TI Uncoupling protein homologs: Emerging views of physiological function, AU Adams, Sean H. (1)
CS (1) Department of Endocrinology, Genentech, Incorporated, South San Francisco, CA, 94080 USA
SO Journal of Nutrition, (April, 2000) Vol. 130, No. 4, pp. 711-714.
ISSN: 0022-3166.

DT Article
LA English
SL English
AB The widespread occurrence of excess weight and related diseases demands that efforts be made to understand energy expenditure from the gene to the whole animal. For some time, it has been understood that mitochondrial oxidation of fuels generates an electrochemical gradient via outward pumping of protons by the electron transport chain. ATP production via F1F0 ATP synthase is then facilitated by the inward flux of protons down the gradient. There is a growing appreciation that a significant portion of the metabolic rate of endotherms is attributable to countering

"proton leak" (uncoupling), wherein a flux of protons down the electrochemical gradient generates heat independently of ATP production. Proton leak is especially apparent in thermogenic brown adipose tissue, which expresses a tissue-specific uncoupling protein (UCP1). The recent discovery of widely expressed putative UCP1 homologs (UCP2, UCP3, UCP4, UCP5/brain mitochondrial carrier protein-1 (BMCP1)) raised the possibility that innate proton leak and metabolic rate are regulated by UCP1-like proteins. On the basis of current published data, one may not exclude the possibility that UCP homologs influence metabolic rate.

L4 ANSWER 5 OF 8 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
AN 2000382329 EMBASE
TI Appetite regulation and uncoupling proteins!
TI APPTITREGELERING OCH UKKOPPLANDE PROTEINER.
AU Erlanson-Albertsson C.
CS Dr. C. Erlanson-Albertsson, Inst. for Cell och Molekylarbiologi, Lunds Universitet, P.O. Box 94, S-221 00 Lund, Sweden, charlotte.erlanson-albertsson@medkem.lu.se
SO Scandinavian Journal of Nutrition/Hairingsforskning, (2000) 44(3) 108-110, Refs: 31
ISSN: 1102-6480 CODEN: SJNUIE
CJ Sweden
DT Journal, General Review
FS 002 Physiology
U29 Clinical Biochemistry
048 Gastroenterology
LA Swedish
SL English
AB Appetite regulation and feeding behavior is critical for survival. In understanding the mechanism for appetite regulation various theories have been postulated. One of these, the thermodynamic theory, postulates that appetite regulation is tightly linked to body temperature, a rise in body temperature causing the termination of a meal. With the discovery of a family of uncoupling proteins, UCP 1-UCP5, the production of heat during feeding has attracted new interest. Special interest has been focused on UCP2, which is produced in the gastrointestinal tract.

L4 ANSWER 6 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
AN 2000253107 BIOSIS
DN PRE200000253107
TI Specific sequence motifs of mitochondrial uncoupling proteins, AU Jezek, Petr (1); Urbanikova, Eva
CS (1) Department of Membrane Transport Biophysics, Institute of Physiology, Academy of Sciences of the Czech Republic, Vitevska 1083, CZ 14220, Prague
Czech Republic
SO IUBMB Life, (Jan., 2000) Vol. 49, No. 1, pp. 63-70, print.
ISSN: 1521-6543.

DT Article
LA English
SL English
AB We have searched for the exclusivity of common sequence motifs of the mitochondrial uncoupling proteins (UCP1, UCP2, UCP3, UCP4, BMCP1, and plant UCP (PUMIP)) within the gene family of mitochondrial anion carrier proteins. The UCP-specific sequences, "UCP signatures", were found in the first, second, and fourth alpha-helices. First: Ala/Ser-Cys/Thr/n-Phe-
Ala/Gly(-)-n-Phe-n-Cys-Thr-Phe/n, second: Gly/Ala-Ile/Ileu-GlnX(+)-NH-
n-Cys-Ser/n-phX-n-Ser-OH/Gly-n-(+)-Ile/Met-Gly/Val-n-Trp/n, fourth:
n-Pro-Asn/Thr-n-X(+)-Asn/Ser/Ile-n-Ile/Ileu-n-Asn/Val-Cys/n-n-Tr(-)-n-
Phe/n-Pro-CH/Val (n, nonpolar, phX, aromatic, +, -, charged residue). The second and part of the third signature are also present in the yeast dicarboxylate transporter. The UCP signature excluding BMCP1 was also found in the second matrix segment (+)-Pro(Ileu-Leu/Ileu)(+)-phX-X-
Gly/Ser-Thr/n-X-NH(-)-Ala-ph. These UCP signatures are thought to be involved in fatty acid anion binding and translocation.

L4 ANSWER 7 OF 8 MEDLINE
AN 2000117196 MEDLINE
DN 20171796 PubMed ID: 10653471
TI Uncoupling protein--a useful energy dissipator, AU Klingenberg M
CS Institute of Physical Biochemistry, University of Munich, Germany
SO JOURNAL OF BIOENERGETICS AND BIOMEMBRANES, (1999 Oct) 31 (6) 419-30. Ref: 80
Journal code: HIO, 7701859, ISSN: 0145-479X

CJ United States
DT Journal, Article, (JOURNAL ARTICLE)
General Review, (REVIEW)
(REVIEW, TUTORIAL)

LA English
FS Priority Journals
EM 200003
ED Entered STN: 20000320
Last Updated on STN: 20000320
Entered Medline: 20000306

AB The structure/function relationship in the uncoupling proteins (UCP) is reviewed, stressing UCP from brown adipose tissue (UCP1) since, so far, nearly no biochemistry is known for the UCP variants UCP2, UCP3, and UCP4.
The transport for H+ and Cl- and its dependence on fatty acids in reconstituted vesicles is described. The inhibition and binding of nucleotides to UCP1, in particular, the pH dependence and two-stage binding are analyzed. A model for the role of fatty acid in H+ transport is shown. The role of specific residues in UCP1 is analyzed by directed mutagenesis in a yeast expression system. The different regulation by the cellular energy potential of UCP1 versus UCP3 is discussed.

L4 ANSWER 8 OF 8 MEDLINE
AN 1999148824 MEDLINE
DN 99148824 PubMed ID: 10025957
TI UCP4, a novel brain-specific mitochondrial protein that reduces membrane potential in mammalian cells, AU Mao W; Yu X X; Zhong A; Li W; Brush J; Sherwood S W; Adams S H; Pan G
CS Department of Endocrinology, Genentech, Inc., South San Francisco, CA 94080, USA.
SO FEBS LETTERS, (1999 Jan 29) 443 (3) 326-30
Journal code: EUH, 0156157, ISSN: 0014-5793.
CJ Netherlands
DT Journal, Article, (JOURNAL ARTICLE)
LA English
FS Priority Journals
OS GENBANK-AF110532
EM 199903
ED Entered STN: 19990324
Last Updated on STN: 20000303
Entered Medline: 19990311

AB Uncoupling proteins (UCPs) are a family of mitochondrial transporter proteins that have been implicated in thermoregulatory heat production and maintenance of the basal metabolic rate. We have identified and partially characterized a novel member of the human uncoupling protein family, termed uncoupling protein-4 (UCP4). Protein sequence analyses showed that UCP4 is most related to UCP3 and possesses

features characteristic of mitochondrial transporter proteins. Unlike other known UCPs, UCP4 transcripts are exclusively expressed in both fetal and adult brain tissues. UCP4 maps to human chromosome 6p11.2-q12. Consistent with its potential role as an uncoupling protein, UCP4 is localized to the mitochondria and its ecotopic expression in mammalian cells reduces mitochondrial membrane potential. These findings suggest that UCP4 may be involved in thermoregulatory heat production and metabolism in the brain.

=> logoff

ALL L# QUERIES AND ANSWER SETS ARE DELETED AT LOGOFF
LOGOFF? (Y/N)HOLD Y

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
FULL ESTIMATED COST	23 49	SESSION	23 64

STN INTERNATIONAL LOGOFF AT 12:41:31 ON 03 MAY 2001

